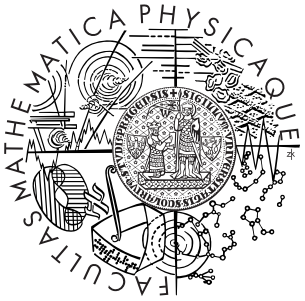
Charles University in Prague
Faculty of Mathematics and Physics

MASTER THESIS



Ondřej Čertík

Calculation of electron structure in the framework of DFT in real space

Institute of Theoretical Physics

Supervisor: RNDr. Jiří Vackář, CSc. Institute of Physics,
Academy of Sciences, Czech Republic

Field of study: Theoretical Physics

2008

I would like to thank my supervisor RNDr. Jiří Vackář, CSc. for his time spent in countless discussions enlightening me on a lot of issues in quantum mechanics and helping with this work. I thank RNDr. Ondřej Šipr, CSc. for explaining me a number of theoretical problems and general advice. I am deeply indebted to RNDr. Antonín Fejfar, CSc. and Ing. Jiří Plešek, CSc. for moral support during my university years and in the course of writing this diploma thesis. I also wish to thank Nissy Nevil for proofreading and her amiable encouragement. Finally, I thank Ing. Robert Cimrman, Ph.D. and his sfepy (a finite element package) and RNDr. Pavel Šolín, Ph.D. for the feedback and help with finite elements.

I declare that I wrote the thesis by myself and listed all used sources. I agree with making the thesis publicly available.

Prague, August 5, 2008 Ondřej Čertík

Title: Calculation of electron structure in the framework of DFT in real space
Author: Ondřej Čertík
Department: Institute of Theoretical Physics
Supervisor: RNDr. Jiří Vackář, CSc.
Supervisor's e-mail address: vackar@fzu.cz

Abstract: In the present work we study ab-initio electronic structure calculations in real space using density functional theory (DFT), finite elements and pseudopotentials. We summarize the theory and full ab-initio derivation of all equations in finite elements, density functional theory and pseudopotentials, then we explain how our program works and we show results for spherically symmetric potentials in relativistic and nonrelativistic DFT and for 2D and 3D Schrödinger equation for symmetric and non-symmetric potentials.

Keywords: Finite element method, Density functional theory

Název práce: Výpočty elektronových struktur v rámci DFT v reálném prostoru
Autor: Ondřej Čertík
Katedra (ústav): Ústav teoretické fyziky
Vedoucí bakalářské práce: RNDr. Jiří Vackář, CSc.
e-mail vedoucího: vackar@fzu.cz

Abstrakt: V předložené práci studujeme ab-initio výpočty elektronových struktur v rámci teorie funkcionálu hustoty (DFT), konečných prvků a pseudopotenciálů. Shrneme teorii a odvození všech potřebných vztahů v metodě konečných prvků, teorii funkcionálu hustoty a pseudopotenciálů z prvních principů, pak vysvětlíme jak náš program funguje a ukážeme výsledky pro sféricky symetrický potenciál v relativistické i nerelativistické verzi DFT a 2D a 3D Schrödingerovu rovnici pro symetrické a nesymetrické potenciály.

Klíčová slova: metoda konečných prvků, teorie funkcionálu hustoty

Contents

Chapter 1

Introduction

Our ultimate goal is to develop a new code for real-space electronic structure calculations within density functional theory combining the advantages of ab-initio pseudopotentials and finite element method for calculating electronic states and related quantities of complex non-periodic systems. This complex task that people started to tackle quite recently [15, 16, 7, 6] entails a series of new partial problems to be solved.

We describe our motivation in detail in Chapter 2.

The related theory and full ab-initio derivation of all the equations in finite elements, density functional theory and pseudopotentials are summarized in Chapter 3. This chapter is split into parts describing the basic ingredients of our recently developed method: the subsections 3.1. – 3.4. deal respectively, with Dirac and Schrödinger equations, with the density functional theory and Kohn-Sham equations, with ab-initio pseudopotentials including their separable form, and with the finite element method.

Chapter 4 summarizes the results obtained so far within the present thesis, explaining how the pieces of our code works and showing the examples of finite-element method applications for spherically symmetric potentials in relativistic and nonrelativistic DFT and for 2D and 3D Schrödinger equation for symmetric and non-symmetric potentials.

Finally the appendix contains some useful derivations and explanations like delta functions, variations, functional derivatives, Dirac notation and more.

Chapter 2

Motivation, Tasks

2.1 Formulation of the Schrödinger Equation for Finite Element Method

The main objective of the present work is to be able to calculate large electron structure problems using density functional theory, pseudopotentials and finite element method.

In order to get there, we need to decompose the whole problem into smaller tasks, build robust solvers for smaller tasks and then use them like a child assembly game. The two most important steps are the following.

One ingredient is a working density functional theory solver. This will be described in a next section and results presented in the Result's section.

Another ingredient is a robust finite element solver for a one particle Schrödinger equation. We wrote a very general solver that works both in 2D and 3D and can solve the eigenproblem for any potential accurately and reliably. We calculated many well known solutions in 2D and 3D (for spherically symmetric problems) and checked with analytical solutions to be sure the solver works well. We then also calculated a couple of nonsymmetric problems in 2D to show the usefulness of our solver – those are unique results that are quite difficult to get using other means – our solver solves symmetric and nonsymmetric problems with ease.

2.1.1 Weak Formulation

See the section 3.4 for the introduction to the finite element method. Here we just say that the FEM solves the so called weak formulation

$$\int \frac{\hbar^2}{2m}\nabla\psi\cdot\nabla v\,\mathrm{d}V+\int vV\psi\,\mathrm{d}V=\int E\psi v\,\mathrm{d}V+\oint\frac{\hbar^2}{2m}\frac{\mathrm{d}\psi}{\mathrm{d}n}v\,\mathrm{d}S\,, \qquad (2.1)$$

The problem reads: find a function ψ such that (3.41) holds for every v.

We choose a basis ϕ_i and substitute ϕ_i for v and expand $\psi=\sum q_j\phi_j$

$$\left(\int\frac{\hbar^2}{2m}\nabla\phi_j\cdot\nabla\phi_i\,\mathrm{d}V+\int\phi_iV\phi_j\,\mathrm{d}V\right)q_j=\left(\int E\phi_j\phi_i\,\mathrm{d}V\right)q_j+\oint\frac{\hbar^2}{2m}\frac{\mathrm{d}\psi}{\mathrm{d}n}\phi_i\,\mathrm{d}S\,, \qquad (2.2)$$

which can be written in a matrix form

$$\left(K_{ij}+V_{ij}\right)q_j=EM_{ij}q_j+F_i\,,$$

where

$$\begin{aligned}
V_{ij} &= \int\phi_iV\phi_j\,\mathrm{d}V\,,\\
M_{ij} &= \int\phi_i\phi_j\,\mathrm{d}V\,,\\
K_{ij} &= \frac{\hbar^2}{2m}\int\nabla\phi_i\cdot\nabla\phi_j\,\mathrm{d}V\,,\\
F_i &= \frac{\hbar^2}{2m}\oint\frac{\mathrm{d}\psi}{\mathrm{d}n}\phi_i\,\mathrm{d}S\,.
\end{aligned}$$

Usually we set $F_i=0$.

This is a generalized eigenvalue problem, that needs to be solved in our program. For more details, see the section 3.4.

2.2 Density Functional Theory Formulation for Spherically Symmetric Problems

One of the approaches to calculate large electronic structure problems is density functional theory, which allows us to calculate a ground state charge

density, from which one can calculate a lot of interesting physical properties of the system.

See the section 3.1 for thorough derivation of all equations in DFT and explanation how everything works. It is useful to note here, that for spherically symmetric problems (for example atoms), all that is needed is to have a radial Schrödinger or Dirac solver and that's exactly what we did to test the DFT self-consistency cycle, Poisson's equation, exchange and correlation potential and other things, see the section 3.1 for theory and section 4.2 for results of calculations for Boron and Lead.

2.3 Pseudopotentials

Ab-inito pseudopotentials [17] considerably improve the efficiency of electronic structure calculations of complex systems by separating (and treating in different way) the core and valence electrons, however we define them. Using the modern, ab-initio environment reflecting pseudopotentials [20] in fact does not represent any additional approximation except for linearizing the density functional approach with respect to one-electron energies, which is the common approximation in all standard existing DFT methods. For applying pseudopotentials within the finite-element basis set, using the separable form is necessary.

In combination with the finite element method, the pseudopotentials improve the efficiency of complex system calculations in essential and extremely significant way as a consequence of several mechanisms:

- By treating the core electrons separately in the pseudopotential generation process, the need to include them into the most computationaly expensive DFT selfconsistency iterative process is eliminated. The fact that the core electrons do not participate in this process significantly reduces the number of electronic states and, by that, reduces the number of eigenvectors that have to be found in each step iteratively repeated until the eigenvectors satisfy the DFT selfconsistency condition.

- The relativistic effect can be taken into account just for the core electrons with high binding energies. The separate spin-up and spin-down pseudopotentials can be generated in the cases where taking into account the spin-orbit coupling is desirable for valence electrons, elimi-

nating the need to solve Dirac equation for valence electrons by finite-element method, which would be extremely complicated task.

- Eliminating the high binding energies of core electrons from the DFT selfconsistency iterative process reduces the numerical problems and improves numerical precision.

- Pseudopotentials eliminate high potential gradients and Coulomb potential singularities, which significantly reduces the density of finite-element mesh that is necessary for achieving well converged accurate results. By that, the efficiency can be significantly improved, particularly in the tasks like molecular dynamics calculations.

Chapter 3

Methods Used

3.1 Radial Schrödinger and Dirac Equations

For the general treatment, together with derivation of all the different forms of radial Dirac equations used in the literature, see the author's bachelor thesis [23]. Here we just summarize the results.

3.1.1 Radial Schrödinger Equation

We have a spherically symmetric potential energy

$$V(\mathbf{x}) = V(r)\,.$$

State with a given square of an angular momentum (eigenvalue $l(l+1)$) and its z component (eigenvalue m) is described by the wave function

$$\psi_{nlm}(\mathbf{x}) = R_{nl}(r)\, Y_{lm}\left(\frac{\mathbf{x}}{r}\right)\,, \tag{3.1}$$

where $R_{nl}(r)$ obeys the equation [5] (eq. 2.400)

$$R''_{nl} + \frac{2}{r}R'_{nl} + \frac{2M}{\hbar^2}(E-V)R_{nl} - \frac{l(l+1)}{r^2}R_{nl} = 0\,. \tag{3.2}$$

This is called the radial Schrödinger equation which we want to solve numerically.

3.1.2 Numerical integration for a given energy

Equation (3.2) is the linear ordinary differential equation of the second order, so the general solution is a linear combination of two independent solutions. Normally, the 2 constants are determined from initial and/or boundary conditions. In our case, however, we don't have any other condition besides being interested in solutions that we can integrate on the interval $(0, \infty)$ (and which are normalizable), more exactly we want $R \in L^2$ and $\int_0^\infty r^2 R^2 \,\mathrm{d}r = 1$.

It can be easily shown by a direct substitution, that there are only two asymptotic behaviors near the origin: r^l and r^{-l-1}. We are interested in quadratic integrable solutions only, so we are left with r^l and only one integration constant, which we calculate from a normalization. This determines the solution uniquely.

All the integration algorithms need to evaluate R'', which is a problem at the origin, where all the terms in the equation are infinite, although their sum is finite. We thus start to integrate the equation at some small r_0 (for example $r_0 = 10^{-10}$ a.u.), where all the terms in the equation are finite. If we find the initial conditions $R(r_0)$ and $R'(r_0)$, the solution is then fully determined.

If r_0 is sufficiently small, we can set $R(r_0) = r_0^l$ and $R'(r_0) = l r_0^{l-1}$. In the case $l = 0$ we need to set $R(r_0) = 1$ and $R'(r_0) = -\frac{1}{a}$, where a is the Bohr radius, see the next section for more details.

So when somebody gives us l and E, we are now able to compute the solution up to the the multiplicative constant that is later determined from a normalization. We used the fourth-order Runge-Kutta method that proved very suitable for this problem.

3.1.3 Asymptotic behavior

The asymptotic behavior is important for the integration routine to find the correct solution for a given E. It is well known, that the first term of the Taylor series of the solution is r^l, independent of the potential [5] (eq. 2.408). This is enough information to find the correct solution for $l > 0$ because the only thing we need to know is the value of the wave function and its derivative near the origin, which is effectively r_0^l and $l r_0^{l-1}$ for some small r_0. The problem is with $l = 0$, where the derivative cannot be calculated just from l and r_0.

The asymptotic behavior for $l = 0$ depends on the potential V, so we

need to take into account it's properties. We assume V to be of a form:

$$V = -\frac{Z}{r} + v_0 + v_1 r + O(r^2)\,,$$

It can be shown, that the solution is then

$$R = a_0(1 - \frac{r}{a} + O(r^2))\,,$$

where $a = \frac{\hbar^2}{ZM}$ is the Bohr radius and a_0 is a normalization constant. So the initial condition for the integration for $l = 0$ is $R(r_0) = 1$ and $R'(r_0) = -\frac{1}{a}$.

3.1.4 Dirac Equation

The Dirac equation for one particle is [19, 22]:

$$H\psi = W\psi\,, \tag{3.3}$$

$$H = c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2 + V(r)\mathbb{1}\,,$$

where ψ is a four component vector:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix}\,, \qquad \psi_A = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix},\ \psi_B = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}$$

and $\boldsymbol{\alpha}$, β are 4×4 matrices:

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}\,,$$

$$\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}\,,$$

where the Pauli matrices $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and $\mathbb{1}$ form a basis of all 2×2 Hermitian matrices. To derive a continuity equation, we multiply (3.3) by ψ^* and subtract the conjugate transpose of (3.3) multiplied by ψ:

$$\frac{\partial}{\partial t}(\psi^*\psi) = -\nabla\cdot(c\psi^*\boldsymbol{\alpha}\psi)\,,$$

so we identify the probability and current densities as

$$\rho = \psi^*\psi = \psi_1^*\psi_1 + \psi_2^*\psi_2 + \psi_3^*\psi_3 + \psi_4^*\psi_4\,, \qquad \mathbf{j} = c\psi^*\boldsymbol{\alpha}\psi\,.$$

The normalization of a four-component wave function is then

$$\int \rho\,\mathrm{d}^3x = \int \psi^*\psi\,\mathrm{d}^3x = \int \psi_1^*\psi_1 + \psi_2^*\psi_2 + \psi_3^*\psi_3 + \psi_4^*\psi_4\,\mathrm{d}^3x = 1\,. \tag{3.4}$$

The probability density $\rho(x, y, z)$ is the physical quantity we are interested in, and all the four-component wavefunctions and other formalism is just a way of calculating it. This ρ is also the thing we should compare with the probability density calculated from the solution of Schrödinger equation.

3.1.5 Radial Dirac equation

We search for a basis in the form of spin angular functions:

$$\psi_A = g\chi_\kappa^{j_3}\,, \tag{3.5}$$

$$\psi_B = if\chi_{-\kappa}^{j_3}\,. \tag{3.6}$$

Substituting all of these into (3.3) and some more well-known manipulations one gets:

$$\hbar c \begin{pmatrix} -\frac{\partial f}{\partial r} + \frac{\kappa-1}{r} f \\ \frac{\partial g}{\partial r} + \frac{\kappa+1}{r} g \end{pmatrix} = \begin{pmatrix} (W - V - mc^2)g \\ (W - V + mc^2)f \end{pmatrix}. \tag{3.7}$$

This is the radial Dirac equation. As we shall see in the next section, the equation for g is (with the exception of a few relativistic corrections) identical to the radial Schrödinger equation and f vanishes in the limit $c \to \infty$. For this reason f is called the small (fein, minor) component and g the large (groß, major) component.

The probability density is

$$\rho = \psi^*\psi = \psi_A^*\psi_A + \psi_B^*\psi_B = f^2\chi_{-\kappa}^{j_3*}\chi_{-\kappa}^{j_3} + g^2\chi_\kappa^{j_3*}\chi_\kappa^{j_3}\,,$$

so from the normalization condition (3.4) we get

$$\int \rho\,\mathrm{d}^3x = \int f^2\chi_{-\kappa}^{j_3*}\chi_{-\kappa}^{j_3} + g^2\chi_\kappa^{j_3*}\chi_\kappa^{j_3}\,\mathrm{d}^3x = \int (f^2\chi_{-\kappa}^{j_3*}\chi_{-\kappa}^{j_3} + g^2\chi_\kappa^{j_3*}\chi_\kappa^{j_3})\,r^2\,\mathrm{d}r\mathrm{d}\Omega =$$

$$= \int_0^\infty f^2r^2\,\mathrm{d}r \int \chi_{-\kappa}^{j_3*}\chi_{-\kappa}^{j_3}\,\mathrm{d}\Omega + \int_0^\infty g^2r^2\,\mathrm{d}r \int \chi_\kappa^{j_3*}\chi_\kappa^{j_3}\,\mathrm{d}\Omega = \int_0^\infty r^2(f^2+g^2)\,\mathrm{d}r = 1\,,$$

where we used the normalization of spin-angular functions. Also it can be seen, that the radial probability density is

$$\rho(r) = r^2(f^2 + g^2) \tag{3.8}$$

(i.e., the probability to find the electron between r_1 and r_2 is $\int_{r_1}^{r_2} r^2(f^2 + g^2)\,\mathrm{d}r$). The result of integrating the radial Dirac equation are the two functions f and g, but the physically relevant quantity is the radial probability density (3.8). In the nonrelativistic case, the density is given by

$$\rho(r) = r^2 R^2\,,$$

so the correspondence between the Schrödinger and Dirac equation is $R^2 = f^2 + g^2$.

For numerical stability and robustness, we are not solving the equations in the form (3.7), but a slightly rearranged ones. Let's use Hartree atomic units ($m = \hbar = 1$) and define $E = W - mc^2 = W - c^2$, so that E doesn't contain the electron rest mass energy. Let's make the substitution [13]

$$P_\kappa = r g_\kappa\,,$$

$$Q_\kappa = r f_\kappa$$

and plug all of this into (3.7). After a little manipulation we get:

$$\frac{\mathrm{d}P_\kappa}{\mathrm{d}r} = -\frac{\kappa}{r} P_\kappa + \left[\frac{E - V}{c} + 2c\right] Q_k\,,$$

$$\frac{\mathrm{d}Q_\kappa}{\mathrm{d}r} = \frac{\kappa}{r} Q_\kappa - \frac{1}{c}(E - V) P_k\,, \tag{3.9}$$

which can be found in [22] (eq. 8.12 and 8.13), where they have one c hidden in $Q_\kappa = cr f_\kappa$ and use Rydberg atomic units, so they have 1 instead of 2 in the square bracket. It can be found in [1] as well, they use Hartree atomic units, but have a different notation $G_\kappa \equiv P_\kappa$ and $F_\kappa \equiv Q_\kappa$, also they made a substitution $c = \frac{1}{\alpha}$.

3.1.6 Asymptotic behavior

We calculate the functions f_κ and g_κ in a similar way as we calculated R for the Schrödinger equation, thus we need the asymptotic behavior at the

origin. The potential can always be treated as $V = 1/r + \cdots$ and in this case it can be shown [22], that the asymptotic is

$$P_\kappa = rg_\kappa = r^\beta \,,$$

$$Q_\kappa = rf_\kappa = r^{\beta-1}\frac{\beta+\kappa}{\frac{E-V}{c}+2c} \,,$$

where

$$\beta = \sqrt{\kappa^2 - \left(\frac{Z}{c}\right)^2} \,, \tag{3.10}$$

or, if we write it explicitly, for $j = l + \frac{1}{2}$

$$\beta^+ = \sqrt{(-l-1)^2 - \left(\frac{Z}{c}\right)^2}$$

and $j = l - \frac{1}{2}$

$$\beta^- = \sqrt{l^2 - \left(\frac{Z}{c}\right)^2} \,.$$

In the semirelativistic case (which is an approximation — we neglect the spin-orbit coupling term) we choose

$$\beta = \sqrt{\tfrac{1}{2}(|\beta^+|^2 + |\beta^-|^2)} = \sqrt{l^2 + l + \tfrac{1}{2} - \left(\frac{Z}{c}\right)^2} \,.$$

It should be noted that in the literature we can find other types of asymptotic behavior for the semirelativistic case, it's just a question of the used approximation. One can hardly say that some of them are correct and another is not since the semirelativistic (sometimes denoted as scalar-relativstic) approximation itself is not correct, it's just an approximation.

It follows from (3.10) that for $j = l + \frac{1}{2}$ the radial Dirac equation completely becomes the radial Schrödinger equation in the limit $c \to \infty$ (and gives exactly the same solutions):

$$P_\kappa = rg_\kappa \to r^{l+1} \,,$$

$$Q_\kappa = rf_\kappa \to 0 \,.$$

For $j = l - \frac{1}{2}$ however, we get a wrong asymptotic: we get a radial Schrödinger equation for l, but the asymptotic for $l - 1$.

3.1.7 Eigenproblem

In the previous sections, we learned how to calculate the solution of both the radial Schrödinger and Dirac equations for a given E. For most of the energies, however, the solution for $r \to \infty$ exponentially diverges to $\pm\infty$. Only for the energies equal to eigenvalues, the solution tends exponentially to zero for $r \to \infty$. The spectrum for bounded states is discrete, so we label the energies by n, starting from 1.

We want to find the eigenvalue and eigenfunction for a given n and l (and a spin in the relativistic case). The algorithm is the same for both nonrelativistic and relativistic case and is based on two facts, first that the number of nodes (ie. the number of intersections with the x axis, not counting the one at the origin and in the infinity) of R_{nl} and g_κ is $n-l-1$ and second that the solution must tend to zero at infinity.

We calculate the solution for some (random) energy E_0, using the procedure described above. Then we count the number of nodes (for diverging solutions, we don't count the last one) and check, if the solution is approaching the zero from top or bottom in the infinity. From the number of nodes and the direction it is approaching the zero it can be determined whether the energy E_0 is below or above the eigenvalue E belonging to a given n and l. The rest is simple, we find two energies, one below E, one above E and by bisecting the interval we calculate E with any precision we want.

There are a few technical numerical problems that are unimportant from the theoretical point of view, but that need to be solved if one attempts to actually implement this algorithm. One of them is that when the algorithm (described in the previous paragraph) finishes, because the energy interval is sufficiently small, it doesn't mean the solution is near zero for the biggest r of our grid. Remember, the solution goes exponentially to $\pm\infty$ for every E except the eigenvalues and because we never find the exact eigenvalue, the solution will (at some point) diverge from zero.

Possible solution that we have employed is as follows: when the algorithm finishes we find the last minimum (which is always near zero) and trim the solution behind it (set it to zero).

The second rather technical problem is how to choose the initial interval of energies so that the eigenvalue lies inside the interval. We use some default values that work for atomic calculations, while allowing the user to override it if needed.

3.2 Density Functional Theory

3.2.1 Introduction

Good books about DFT are [3] and [14], but they both contain much more topics which we don't need and some topics are missing in each of them, so this chapter gives a self-contained explanation of all one has to know about a many body quantum mechanics and the DFT in order to be able to do DFT calculations.

3.2.2 Many Body Schrödinger Equation

We use the Born-Oppenheimer approximation, which says that the nuclei of the treated atoms are seen as fixed. A stationary electronic state (for N electrons) is then described by a wave function $\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r}_N)$ fulfilling the many-body Schrödinger equation

$$\hat{H}\,|\Psi\rangle = (\hat{T} + \hat{U} + \hat{V})\,|\Psi\rangle = E\,|\Psi\rangle$$

where

$$\hat{T} = \sum_{i}^{N} -\frac{1}{2}\nabla_i^2$$

is the kinetic term,

$$\hat{U} = \sum_{i<j} U(\mathbf{r_i}, \mathbf{r_j}) = \frac{1}{2}\sum_{i,j} U(\mathbf{r_i}, \mathbf{r_j})$$

$$U(\mathbf{r_i}, \mathbf{r_j}) = U(\mathbf{r_j}, \mathbf{r_i}) = \frac{1}{|\mathbf{r_i} - \mathbf{r_j}|}$$

is the electron-electron interaction term and

$$\hat{V} = \sum_{i}^{N} v(\mathbf{r_i})$$

$$v(\mathbf{r_i}) = \sum_{k} -\frac{Z_k}{|\mathbf{r_i} - \mathbf{R_k}|}$$

is the interaction term between electrons and nuclei, where R_k are positions of nuclei and Z_k the number of nucleons in each nucleus (we are using atomic

units). So for one atomic calculation with the atom nucleus in the origin, we have just $v(\mathbf{r_i}) = -\frac{Z}{|\mathbf{r_i}|}$.

$|\Psi|^2 = \Psi^*\Psi$ gives the probability density of measuring the first electron at the position $\mathbf{r_1}$, the second at $\mathbf{r_2}$, ... and the Nth electron at the position $\mathbf{r}_N$. The normalization is such that $\int |\Phi|^2 \mathrm{d}^3r_1 \mathrm{d}^3r_2 \ldots \mathrm{d}^3r_N = 1$. The Ψ is anti-symmetric, i.e. $\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r}_N) = -\Psi(\mathbf{r_2}, \mathbf{r_1}, \cdots, \mathbf{r}_N) = -\Psi(\mathbf{r_1}, \mathbf{r}_N, \cdots, \mathbf{r_2})$ etc.

Integrating $|\Psi|^2$ over the first $N-1$ electrons is the probability density that the Nth electron is at the position $\mathbf{r}_N$. Thus the probability density $n(\mathbf{r})$ that any of the N electrons (i.e the first, or the second, or the third, ..., or the Nth) is at the position $\mathbf{r}$ is called the particle (or charge or electron) density and is therefore given by:

$$n(\mathbf{r}) = \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N)\Psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N)\, \mathrm{d}^3r_2\, \mathrm{d}^3r_3 \cdots \mathrm{d}^3r_N +$$

$$+ \int \Psi^*(\mathbf{r}_1, \mathbf{r}, \cdots, \mathbf{r}_N)\Psi(\mathbf{r}_1, \mathbf{r}, \cdots, \mathbf{r}_N)\, \mathrm{d}^3r_1\, \mathrm{d}^3r_3 \cdots \mathrm{d}^3r_N + \cdots$$

$$+ \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r})\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r})\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2\, \mathrm{d}^3r_3 \cdots \mathrm{d}^3r_{N-1} =$$

$$= \int (\delta(\mathbf{r} - \mathbf{r}_1) + \delta(\mathbf{r} - \mathbf{r}_2) + \cdots + \delta(\mathbf{r} - \mathbf{r}_N))$$

$$\Psi^*(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2\, \mathrm{d}^3r_3 \cdots \mathrm{d}^3r_N =$$

$$= \sum_{i=1}^{N} \int \langle \Psi | \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N \rangle\, \delta(\mathbf{r} - \mathbf{r}_i)\, \langle \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N | \Psi \rangle\ \mathrm{d}^3r_1\, \mathrm{d}^3r_2\, \mathrm{d}^3r_3 \cdots \mathrm{d}^3r_N =$$

$$= N \int \langle \Psi | \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N \rangle\, \delta(\mathbf{r} - \mathbf{r}_1)\, \langle \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N | \Psi \rangle\ \mathrm{d}^3r_1\, \mathrm{d}^3r_2\, \mathrm{d}^3r_3 \cdots \mathrm{d}^3r_N =$$

$$= N \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N)\Psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N)\, \mathrm{d}^3r_2\, \mathrm{d}^3r_3 \cdots \mathrm{d}^3r_N \qquad (3.11)$$

Thus $\int_\Omega n(\mathbf{r})\, \mathrm{d}^3r$ gives the number of particles (and also the amount of charge) in the region of integration Ω. Obviously $\int n(\mathbf{r})\, \mathrm{d}^3r = N$.

The energy of the system is given by

$$E = \langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi | \hat{T} | \Psi \rangle + \langle \Psi | \hat{U} | \Psi \rangle + \langle \Psi | \hat{V} | \Psi \rangle = T + U + V \qquad (3.12)$$

where

$$T = \langle\Psi|\hat{T}|\Psi\rangle = \sum_i^N \int \Psi^*(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})(-\tfrac{1}{2}\nabla_i^2)\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2\cdots\mathrm{d}^3r_N$$

$$U = \langle\Psi|\hat{U}|\Psi\rangle$$

$$V = \langle\Psi|\hat{V}|\Psi\rangle = \sum_i^N \int \Psi^*(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})v(\mathbf{r_i})\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2\cdots\mathrm{d}^3r_N =$$

$$= \sum_i^N \int \Psi^*(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})v(\mathbf{r_1})\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2\cdots\mathrm{d}^3r_N =$$

$$= N \int \Psi^*(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})v(\mathbf{r_1})\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2\cdots\mathrm{d}^3r_N =$$

$$= \int v(\mathbf{r})n(\mathbf{r})\mathrm{d}^3r = V[n] \tag{3.13}$$

It needs to be stressed, that E generally is **not** a functional of n alone, only the $V[n]$ is. In the next section we show however, that if the $|\Psi\rangle$ is a ground state (of any system), then E becomes a functional of n.

3.2.3 The Hohenberg-Kohn Theorem

The Schrödinger equation gives the map

$$C : V \to \Psi$$

where Ψ is the ground state. C is bijective (one-to-one correspondence), because to every V we can compute the corresponding Ψ from Schrödinger equation and two different V and V' (differing by more than a constant) give two different Ψ, because if V and V' gave the same Ψ, then by substracting

$$\hat{H}\,|\Psi\rangle = E_{gs}\,|\Psi\rangle$$

from

$$\hat{H}'\,|\Psi\rangle = (\hat{H} - \hat{V} + \hat{V}')\,|\Psi\rangle = E'_{gs}\,|\Psi\rangle$$

we would get $V - V' = E - E'$, which is a contradiction with the assumption that V and V' differ by more than a constant.

Similarly, from the ground state wavefunction Ψ we can compute the charge density n giving rise to the map

$$D : \Psi \to n$$

which is also bijective, because to every Ψ we can compute n from (3.11) and two different Ψ and Ψ' give two different n and n', because different Ψ and Ψ' give

$$E_{gs} = \langle \Psi | \hat{H} | \Psi \rangle < \langle \Psi' | \hat{H} | \Psi' \rangle = \langle \Psi' | \hat{H}' + \hat{V} - \hat{V}' | \Psi' \rangle = E'_{gs} + \int n'(\mathbf{r})(v(\mathbf{r}) - v'(\mathbf{r}))\, \mathrm{d}^3 r$$

$$E'_{gs} = \langle \Psi' | \hat{H}' | \Psi' \rangle < \langle \Psi | \hat{H}' | \Psi \rangle = \langle \Psi | \hat{H} + \hat{V}' - \hat{V} | \Psi \rangle = E_{gs} + \int n(\mathbf{r})(v'(\mathbf{r}) - v(\mathbf{r}))\, \mathrm{d}^3 r$$

adding these two inequalities together gives

$$0 < \int n'(\mathbf{r})(v(\mathbf{r}) - v'(\mathbf{r}))\, \mathrm{d}^3 r + \int n(\mathbf{r})(v'(\mathbf{r}) - v(\mathbf{r}))\, \mathrm{d}^3 r = \int (n(\mathbf{r}) - n'(\mathbf{r}))(v'(\mathbf{r}) - v(\mathbf{r}))\, \mathrm{d}^3 r$$

which for $n = n'$ gives $0 < 0$, which is nonsense, so $n \neq n'$.

So we have proved that for a given ground state density $n_0(\mathbf{r})$ (generated by a potential $\hat{V}_0$) it is possible to calculate the corresponding ground state wavefunction $\Psi_0(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})$, in other words, Ψ_0 is a unique functional of n_0:

$$\Psi_0 = \Psi_0[n_0]$$

so the ground state energy E_0 is also a functional of n_0

$$E_0 = \langle \Psi_0[n_0] | \hat{T} + \hat{U} + \hat{V}_0 | \Psi_0[n_0] \rangle = E[n_0]$$

We define an energy functional

$$E_{v_0}[n] = \langle \Psi[n] | \hat{T} + \hat{U} + \hat{V}_0 | \Psi[n] \rangle = \langle \Psi[n] | \hat{T} + \hat{U} | \Psi[n] \rangle + \int v_0(\mathbf{r}) n(\mathbf{r}) \mathrm{d}^3 r \tag{3.14}$$

where $|\Psi[n]\rangle$ is any ground state wavefunction (generated by an arbitrary potential), that is, n is a ground state density belonging to an arbitrary system. E_0 which is generated by the potential V_0 can then be expressed as

$$E_0 = E_{v_0}[n_0]$$

and for $n \neq n_0$ we have (from the Ritz principle)

$$E_0 < E_{v_0}[n]$$

and one has to minimize the functional $E_{v_0}[n]$:

$$E_0 = \min_n E_{v_0}[n] \tag{3.15}$$

The term

$$\langle \Psi[n] | \hat{T} + \hat{U} | \Psi[n] \rangle \equiv F[n]$$

in (3.14) is universal in the sense that it doesn't depend on $\hat{V}_0$. It can be proven [3], that $F[n]$ is a functional of n for degenerated ground states too, so (3.15) stays true as well.

The ground state densities in (3.14) and (3.15) are called **pure-state v-representable** because they are the densities of (possible degenerate) ground state of the Hamiltonian with some local potential $v(\mathbf{r})$. One may ask a question if all possible functions are v-representable (this is called the v-representability problem). The question is relevant, because we need to know which functions to take into account in the minimization process (3.15). Even though not every function is v-representable [3], every density defined on a grid (finite of infinite) which is strictly positive, normalized and consistent with the Pauli principle is ensemble v-representable. Ensemble v-representation is just a simple generalization of the above, for details see [3]. In plain words, we are fine.

The functional $E_{v_0}[n]$ in (3.15) depends on the particle number N, so in order to get n, we need to solve the variational formulation

$$\frac{\delta}{\delta n} \left(E_v[n] - \mu(N) \int n(\mathbf{r}) \mathrm{d}^{\mathbf{3}}\mathbf{r} \right) = 0$$

so

$$\frac{\delta E_v[n]}{\delta n} = \mu(N) \tag{3.16}$$

Let the $n_N(\mathbf{r})$ be the solution of (3.16) with a particle number N and the energy E_N:

$$E_N = E_v[n_N]$$

The Lagrangian multiplier μ is the exact chemical potential of the system

$$\mu(N) = \frac{\partial E_N}{\partial N}$$

becuase

$$E_{N+\epsilon} - E_N = E_v[n_{N+\epsilon}] - E_v[n_N] = \int \frac{\delta E_v}{\delta n} (n_{N+\epsilon} - n_N) \mathrm{d}^3 r =$$

$$= \int \mu(N)(n_{N+\epsilon} - n_N)\mathrm{d}^3 r = \mu(N)(N + \epsilon - N) = \mu(N)\epsilon$$

so

$$\mu(N) = \frac{E_{N+\epsilon} - E_N}{\epsilon} \longrightarrow \frac{\partial E_N}{\partial N}$$

3.2.4 The Kohn-Sham Equations

Consider an auxiliary system of N noninteracting electrons (noninteracting gas):

$$\hat{H}_s = \hat{T} + \hat{V}_s$$

Then the many-body ground state wavefunction can be decomposed into single particle orbitals

$$|\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})\rangle = |\psi_1(\mathbf{r})\rangle\, |\psi_2(\mathbf{r})\rangle \cdots |\psi_N(\mathbf{r})\rangle$$

and

$$E_s[n] = T_s[\{\psi_i[n]\}] + V_s[n]$$

where

$$T_s[n] = \langle \Psi[n]|\hat{T}|\Psi[n]\rangle = \sum_i \langle \psi_i| - \tfrac{1}{2}\nabla^2|\psi_i\rangle$$

$$V_s[n] = \langle \Psi[n]|\hat{V}|\Psi[n]\rangle = \int v_s(\mathbf{r})n(\mathbf{r})\mathrm{d}^3 r$$

From (3.16) we get

$$\mu = \frac{\delta E_s[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \frac{\delta V_s[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + v_s(\mathbf{r}) \tag{3.17}$$

Solution to this equation gives the density n_s.

Now we want to express the energy in (3.12) using T_s and E_H for convenience, where E_H is the classical electrostatic interaction energy of the charge distribution $n(\mathbf{r})$:

$$\nabla^2 V_H = n(\mathbf{r})$$

or equivalently

$$E_H[n] = \tfrac{1}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\mathrm{d}^3 r \mathrm{d}^3 r'$$

$$V_H(\mathbf{r}) = \frac{\delta E_H}{\delta n(\mathbf{r})} = \tfrac{1}{2}\int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\mathrm{d}^3 r' \tag{3.18}$$

So from (3.14) we get

$$E[n] = (T+U)[n] + V[n] = T_s[n] + E_H[n] + (T - T_s + U - E_H)[n] + V[n] =$$

$$= T_s[n] + E_H[n] + E_{xc}[n] + V[n] \tag{3.19}$$

The rest of the energy is denoted by $E_{xc} = U - E_H + T - T_s$ and it is called is the exchange and correlation energy functional. From (3.16)

$$\mu = \frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \frac{\delta E_H[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} + \frac{\delta V[n]}{\delta n(\mathbf{r})}$$

From (3.18) we have

$$\frac{\delta E_H}{\delta n(\mathbf{r})} = V_H(\mathbf{r})$$

from (3.13) we get

$$\frac{\delta V[n]}{\delta n(\mathbf{r})} = v(\mathbf{r})$$

we define

$$\frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = V_{xc}(\mathbf{r}) \tag{3.20}$$

so we arrive at

$$\mu = \frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + v(\mathbf{r}) \tag{3.21}$$

Solution to this equation gives the density n. Comparing (3.21) to (3.17) we see that if we choose

$$v_s \equiv V_H + V_{xc} + v$$

then $n_s(\mathbf{r}) \equiv n(\mathbf{r})$. So we solve the Kohn-Sham equations of this auxiliary non-interacting system

$$(-\tfrac{1}{2}\nabla^2 + v_s(\mathbf{r}))\psi_i(\mathbf{r}) \equiv (-\tfrac{1}{2}\nabla^2 + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + v(\mathbf{r}))\psi_i(\mathbf{r}) = \epsilon_i\psi(\mathbf{r}) \tag{3.22}$$

which yield the orbitals ψ_i that reproduce the density $n(\mathbf{r})$ of the original interacting system

$$n(\mathbf{r}) \equiv n_s(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \tag{3.23}$$

The sum is taken over the lowest N energies. Some of the ψ_i can be degenerated, but it doesn't matter - the index i counts every eigenfunction

including all the degenerated. In plain words, the trick is in realizing, that the ground state energy can be found by minimizing the energy functional (3.14) and in rewriting this functional into the form (3.19), which shows that the interacting system can be treated as a noninteracting one with a special potential.

3.2.5 The XC Term

The exchange and correlation functional

$$E_{xc}[n] = (T+U)[n] - E_H[n] - T_S[n]$$

can always be written in the form

$$E_{xc}[n] = \int n(\mathbf{r}')\epsilon_{xc}(\mathbf{r}';n)\mathrm{d}^3r'$$

where the $\epsilon_{xc}(\mathbf{r}';n)$ is called the xc energy density.

Unfortunately, no one knows $\epsilon_{xc}(\mathbf{r}';n)$ exactly (yet). The most simple approximation is the **local density approximation** (LDA), for which the xc energy density ϵ_{xc} at $\mathbf{r}$ is taken as that of a homogeneous electron gas (the nuclei are replaced by a uniform positively charged background, density $n = \text{const}$) with the same local density:

$$\epsilon_{xc}(\mathbf{r};n) \approx \epsilon_{xc}^{LD}(n(\mathbf{r}))$$

The xc potential V_{xc} defined by (3.20) is then

$$V_{xc}(\mathbf{r};n) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = \epsilon_{xc}(\mathbf{r}';n) + \int n(\mathbf{r}')\frac{\delta\epsilon_{xc}(\mathbf{r}';n)}{\delta n(\mathbf{r})}\mathrm{d}^3r'$$

which in the LDA becomes

$$V_{xc}(\mathbf{r};n) = \epsilon_{xc}^{LD}(n) + n\frac{\mathrm{d}\epsilon_{xc}^{LD}(n)}{\mathrm{d}n} = \frac{\mathrm{d}}{\mathrm{d}n}\left(n\epsilon_{xc}^{LD}(n)\right) = V_{xc}^{LD}(n) \qquad (3.24)$$

The xc energy density ϵ_{xc}^{LD} of the homogeneous gas can be computed exactly[14]:

$$\epsilon_{xc}^{LD}(n) = \epsilon_x^{LD}(n) + \epsilon_c^{LD}(n)$$

where the ϵ_x^{LD} is the electron gas exchange term given by[14]

$$\epsilon_x^{LD}(n) = -\frac{3}{4\pi}(3\pi^2 n)^{\frac{1}{3}}$$

the rest of ϵ_{xc}^{LD} is hidden in $\epsilon_c^{LD}(n)$ for which there doesn't exist an analytic formula, but the correlation energies are known exactly from quantum Monte Carlo (QMC) calculations by Ceperley and Alder[17]. The energies were fitted by Vosko, Wilkes and Nussair (VWN) with $\epsilon_c^{LD}(n)$ and they got accurate results with errors less than 0.05 mRy in ϵ_c^{LD}, which means that $\epsilon_c^{LD}(n)$ is virtually known exactly. VWN result:

$$\epsilon_c^{LD}(n) \approx \frac{A}{2}\left\{\ln\left(\frac{y^2}{Y(y)}\right) + \frac{2b}{Q}\arctan\left(\frac{Q}{2y+b}\right) + \right.$$

$$\left. -\frac{by_0}{Y(y_0)}\left[\ln\left(\frac{(y-y_0)^2}{Y(y)}\right) + \frac{2(b+2y_0)}{Q}\arctan\left(\frac{Q}{2y+b}\right)\right]\right\}$$

where $y = \sqrt{r_s}$, $Y(y) = y^2 + by + c$, $Q = \sqrt{4c - b^2}$, $y_0 = -0.10498$, $b = 3.72744$, $c = 12.93532$, $A = 0.0621814$ and r_s is the electron gas parameter, which gives the mean distance between electrons (in atomic units):

$$r_s = \left(\frac{3}{4\pi n}\right)^{\frac{1}{3}}$$

The xc potential is then computed from (3.24):

$$V_{xc}^{LD} = V_x^{LD} + V_c^{LD}$$

$$V_x^{LD} = -\frac{1}{\pi}(3\pi^2 n)^{\frac{1}{3}}$$

$$V_c^{LD} = \frac{A}{2}\left\{\ln\left(\frac{y^2}{Y(y)}\right) + \frac{2b}{Q}\arctan\left(\frac{Q}{2y+b}\right) + \right.$$

$$\left. -\frac{by_0}{Y(y_0)}\left[\ln\left(\frac{(y-y_0)^2}{Y(y)}\right) + \frac{2(b+2y_0)}{Q}\arctan\left(\frac{Q}{2y+b}\right)\right]\right\} +$$

$$-\frac{A}{6}\frac{c(y-y_0) - by_0 y}{(y-y_0)Y(y)}$$

Some people also use Perdew and Zunger formulas, but they give essentially the same results. The LDA, although very simple, is surprisingly successful. More sophisticated approximations exist, for example the generalized gradient approximation (GGA), which sometimes gives better results than the LDA, but is not perfect either. Other options include orbital-dependent (implicit) density functionals or a linear response type functionals, but this

topic is still evolving. The conclusion is, that the LDA is a good approximation to start with, and only when we are not satisfied, we will have to try some more accurate and modern approximation.

RLDA: Relativistic corrections to the energy-density functional were proposed by MacDonald and Vosko and basically are just a change in $\epsilon_x^{LD}(n) \to \epsilon_x^{LD}(n)R$:

$$R = \left[1 - \frac{3}{2}\left(\frac{\beta\mu - \ln(\beta+\mu)}{\beta^2}\right)^2\right]$$

where

$$\mu = \sqrt{1+\beta^2}$$

and

$$\beta = \frac{(3\pi^2 n)^{\frac{1}{3}}}{c}$$

We also need to calculate these derivatives:

$$\frac{\mathrm{d}R}{\mathrm{d}\beta} = -6\frac{\beta\mu - \ln(\beta+\mu)}{\beta^2}\left(\frac{1}{\mu} - \frac{\beta\mu - \ln(\beta+\mu)}{\beta^3}\right)$$

$$\frac{\mathrm{d}\beta}{\mathrm{d}n} = \frac{\beta}{3n}$$

$$\frac{\mathrm{d}\epsilon_x^{LD}}{\mathrm{d}n} = \frac{\epsilon_x^{LD}}{3n}$$

So

$$V_x^{RLD} = \epsilon_x^{LD}R + n\frac{\mathrm{d}\epsilon_x^{LD}R}{\mathrm{d}n} = \frac{4}{3}\epsilon_x^{LD}R + \frac{1}{3}\epsilon_x^{LD}\frac{\mathrm{d}R}{\mathrm{d}\beta}\beta \tag{3.25}$$

For $c \to \infty$ we get $\beta \to 0$, $R \to 1$ and $V_x^{RLD} \to \frac{4}{3}\epsilon_x^{LD} = V_x^{LD}$ as expected, because

$$\lim_{\beta\to 0}\frac{\beta\sqrt{1+\beta^2} - \ln(\beta+\sqrt{1+\beta^2})}{\beta^2} = 0$$

3.2.6 Iteration to Self-consistency

The V_H and V_{xc} potentials in the Kohn-Sham equations (3.22) depend on the solution n thus the KS equations need to be iterated to obtain a self-consistent density. One can regard the KS procedure as a nonlinear operator $\hat{F}$ which satisfies (at the Mth iteration)

$$n_M^{\mathrm{out}} = \hat{F}n_M$$

and the problem is to find the self-consistent density which satisfies

$$n = \hat{F}n$$

Due to the long-range nature of the Coulomb interaction, a small change in the input density n_M can lead to a relatively large change in the output density $\hat{F}n_M$, thus it is not possible to use the output density itself as the input density for the next iteration, since large unstable charge oscillations arise. Rather it is essential to mix input and output densities in an appropriate manner to obtain a new input density.

The $n(\mathbf{r})$ is in practice defined on some grid, or using coefficients of plane waves, local orbitals or the like, which means that the precise relation

$$\mathbb{1} = \int |\mathbf{r}\rangle \langle \mathbf{r}| \, \mathrm{d}^3 r$$

is changed for

$$\mathbb{1} \approx \sum_i |\mathbf{r}_i\rangle \langle \mathbf{r}_i|$$

in the case of a grid (or some other basis like plane waves can be used instead of $|\mathbf{r}_i\rangle$) and $n(\mathbf{r}) = \langle \mathbf{r}|n\rangle$ is approximated by $n(\mathbf{r}_i) = \langle \mathbf{r}_i|n\rangle$. Let

$$\mathbf{x} = (x_1, x_2, x_3, ...), \quad x_i \equiv n(\mathbf{r}_i) = \langle \mathbf{r}_i|n\rangle$$

and

$$\mathbf{F}(\mathbf{x}_M) \equiv \hat{F}n_M, \quad F_i = (\hat{F}n_M)(\mathbf{r}_i)$$

the self-consistency is reached when $\mathbf{F}(\mathbf{x}) = \mathbf{x}$.

So the problem is in solving the equation

$$\mathbf{F}(\mathbf{x}) = \mathbf{x}$$

where $\mathbf{x}$ denotes a vector in many dimensions (the number of points in the grid). It can also be expressed in the form of the residual $\mathbf{R}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) - \mathbf{x}$ as

$$\mathbf{R}(\mathbf{x}) = 0$$

Almost all of the methods start with approximating

$$\mathbf{R}(\mathbf{x}_{M+1}) - \mathbf{R}(\mathbf{x}_M) \approx \mathbf{J} \cdot (\mathbf{x}_{M+1} - \mathbf{x}_M) \tag{3.26}$$

where the Jacobian

$$J_{ij} = \frac{\partial R_i}{\partial x_j}$$

We want $\mathbf{R}(\mathbf{x}_{M+1}) = 0$, so substituting that into (3.26) we get

$$\mathbf{x}_{M+1} \approx \mathbf{x}_M + \mathbf{J}^{-1} \cdot (\mathbf{R}(\mathbf{x}_{M+1}) - \mathbf{R}(\mathbf{x}_M)) = \mathbf{x}_M - \mathbf{J}^{-1} \cdot \mathbf{R}(\mathbf{x}_M)$$

If we knew the Jacobian exactly, this would be the multidimensional Newton-Raphson method, but we can only make approximations to $\mathbf{J}$ using a sequence of $\mathbf{J}_0$, $\mathbf{J}_1$, $\mathbf{J}_2$, ...:

$$\mathbf{x}_{M+1} = \mathbf{x}_M - \mathbf{J}_M^{-1} \cdot \mathbf{R}(\mathbf{x}_M) \tag{3.27}$$

and the rate of convergence is determined by the quality of the Jacobian. These type of methods are called quasi-Newton-Raphson methods.

The simplest approach is to use the **linear mixing** scheme for which

$$\mathbf{J}_M^{-1} = -\alpha \mathbb{1}$$

so

$$\mathbf{x}_{M+1} = \mathbf{x}_M + \alpha \mathbf{R}(\mathbf{x}_M) = \mathbf{x}_M + \alpha(\mathbf{F}(\mathbf{x}_M) - \mathbf{x}_M)$$

where $0 < \alpha \leq 1$ is the mixing parameter, working value is somewhere around $\alpha = 0.1$ to $\alpha = 0.3$. Unfortunately, this procedure is slow and also we do not explore all the possible densities with this mixing, which means that we don't get the correct density with any accuracy, because we get stuck at a "stiff" situation for which continued iteration does not improve the distance $|\mathbf{R}(\mathbf{x}_M)|$ between input and output densities. On the other hand it's very easy to implement and it works in most cases, although slowly.

Surprisingly very good method is this:

$$\mathbf{J}_M^{-1} = -\mathrm{diag}(\beta_1, \beta_2, \beta_3, \dots) \tag{3.28}$$

start with $\beta_1 = \beta_2 = \beta_3 = \cdots = \alpha$ and at every iteration adjust the parameters β_i according to this very simple algorithm: if $R_i(\mathbf{x}_{M-1})R_i(\mathbf{x}_M) > 0$ then increase β_i by α (if $\beta_i > \alpha_{max}$, set $\beta_i = \alpha_{max}$) otherwise set $\beta_i = \alpha$. In my tests it behaves almost as well as the second Broyden method.

More sophisticated approach is the Broyden update, which updates the $\mathbf{J}$ successively at every iteration. The **first Broyden method** is using this formula:

$$\mathbf{J}_{M+1} = \mathbf{J}_M - \frac{(\Delta\mathbf{R}(\mathbf{x}_M) + \mathbf{J}_M \cdot \Delta\mathbf{x}_M)\Delta\mathbf{x}_M^T}{|\Delta\mathbf{x}_M|^2}$$

which has the disadvantage that we need to compute the inverse Jacobian in (3.27) at every iteration, which is impossible in our case. The **second Broyden method** updates the inverse Jacobian directly using this formula

$$\mathbf{J}_{M+1}^{-1} = \mathbf{J}_M^{-1} + \frac{(\Delta\mathbf{x}_M - \mathbf{J}_M^{-1} \cdot \Delta\mathbf{R}(\mathbf{x}_M))\Delta\mathbf{R}(\mathbf{x}_M)^T}{|\Delta\mathbf{R}(\mathbf{x}_M)|^2} \tag{3.29}$$

starting with the linear mixing:

$$\mathbf{J}_0^{-1} = -\alpha\mathbb{1}$$

It is impossible to store the whole dense matrix of the inverse Jacobian, but fortunately it is not necessary, realizing that the (3.29) has a very simple structure [18]:

$$\mathbf{J}_{M+1}^{-1} = \mathbf{J}_M^{-1} + \mathbf{u}\mathbf{v}^T$$

with

$$\mathbf{u} = \Delta\mathbf{x}_M - \mathbf{J}_M^{-1} \cdot \Delta\mathbf{R}(\mathbf{x}_M) \tag{3.30}$$

$$\mathbf{v} = \frac{\Delta\mathbf{R}(\mathbf{x}_M)}{|\Delta\mathbf{R}(\mathbf{x}_M)|^2}$$

so the whole inverse Jacobian can be written as

$$\mathbf{J}_M^{-1} = -\alpha\mathbb{1} + \mathbf{u}_1\mathbf{v}_1^T + \mathbf{u}_2\mathbf{v}_2^T + \mathbf{u}_3\mathbf{v}_3^T + \cdots$$

and we only need to know how to apply such a Jacobian to an arbitrary vector, which is needed in (3.30) and (3.27):

$$\mathbf{J}_M^{-1} \cdot \mathbf{y} = -\alpha\mathbf{y} + \mathbf{u}_1(\mathbf{v}_1^T\mathbf{y}) + \mathbf{u}_2(\mathbf{v}_2^T\mathbf{y}) + \mathbf{u}_3(\mathbf{v}_3^T\mathbf{y}) + \cdots$$

Thus instead of the whole dense matrix, we only need to save the vectors $\mathbf{u}$ and $\mathbf{v}$ from every iteration.

Vanderbilt and Louie [21] suggested a **modified Broyden method**, which incorporates weights, but Eyert [4] showed that if all the weights are used to tune the iteration process to its fastest convergence, they, in fact, cancel out and the result of the scheme is called by Eyert the **generalized Broyden method**, whose scheme shown by Eyert is exactly the same as for the **Anderson mixing**:

$$\sum_{p=M-k}^{M-1} (1 + \omega_0^2\delta_{pn})\Delta\mathbf{R}(\mathbf{x}_n)^T\Delta\mathbf{R}(\mathbf{x}_p)\gamma_p = \Delta\mathbf{R}(\mathbf{x}_n)^T\mathbf{R}(\mathbf{x}_M)$$

$$\mathbf{x}_{M+1} = \mathbf{x}_M + \beta_M \mathbf{R}(\mathbf{x}_M) - \sum_{p=M-k}^{M-1} \gamma_p (\Delta \mathbf{x}_p + \beta_M \Delta \mathbf{R}(\mathbf{x}_p))$$

which according to Eyert should converge even faster than the second Broyden method, but it doesn't in my own implementation. ω_0 is added just for a numerical stability, good value is $\omega = 0.01$, but it can also be switched off by $\omega_0 = 0$. p is the number of last iterations to use, good value according to Eyert is $p = 5$, β_M shouldn't influence the convergence much for $p = 5$.

The problem with n is that there are two conditions which need to be satisfied

$$n > 0$$

and the normalization

$$\int n(\mathbf{r}) \mathrm{d}^3 r = Z$$

The Newton method converges to the correct norm, but slowly. The condition $n > 0$ however causes great instability. One option could be to use a logistic function like

$$n(r) = \frac{C}{1 + e^{-x(r)}}$$

for sufficiently large C and solve for x, which can be both positive and negative. But more elegant solution is to mix $V_h + V_{xc}$ instead of the densities.

3.2.7 Example: Pb Atom

To illustrate the explained theory, we will show how to calculate the Pb atom. We have $N = 82$ and

$$v(\mathbf{r}_i) = -\frac{82}{|\mathbf{r}_i|}$$

and we need to sum over the lowest 82 eigenvalues in (3.23). One option (the correct one) is to automatically try different "n" and "l" until we are sure we got the lowest 82 energies. But for Pb the combination of "n" and "l" is well-known, it is (first number is n, the letter is l and the number in superscript gives the number of times the particular eigenvalue needs to be taken into account in the sum): $1S^2$, $2S^2$, $2P^6$, $3S^2$, $3P^6$, $3D^{10}$, $4S^2$, $4P^6$, $4D^{10}$, $4F^{14}$, $5S^2$, $5P^6$, $5D^{10}$, $6S^2$, $6P^2$ (notice the 5F and 5G are missing). Together it is 82 eigenvalues. The KS energies for these eigenvalues are:

-2701.6 -466.18 -471.87 -111.45 -108.24 -92.183 -24.498 -22.086 -15.119 -5.6606 -3.9570 -2.9743 -.92718 -.33665 -.14914

3.3 Pseudopotentials

3.3.1 Introduction

Literature about pseudopotentials is unfortunately scattered among many articles, so this section tries to give an overview so that we have everything at one place.

3.3.2 Hermitian Operators in Spherical Symmetry

We show that every Hermitian operator $\hat{V}$ in the spherical symmetric problem ($\hat{V} = R^{-1}\hat{V}R$) can be written in the form

$$\hat{V} = \sum_{lm} |lm\rangle \, \hat{V}_l \, \langle lm| \tag{3.31}$$

where the operator $\hat{V}_l = \langle lm|\hat{V}|lm\rangle$ has matrix elements

$$\langle \rho|\hat{V}_l|\rho'\rangle = \langle lm| \, \langle \rho|\hat{V}|\rho'\rangle \, |lm\rangle = V_l(\rho, \rho')$$

Proof: Matrix elements of a general Hermitian operator $\hat{V}$ are

$$\langle \mathbf{r}|\hat{V}|\varphi\rangle = \int \langle \mathbf{r}|\hat{V}|\mathbf{r}'\rangle \, \langle \mathbf{r}'|\varphi\rangle \, \mathrm{d}^3r' = \int V(\mathbf{r}, \mathbf{r}')\varphi(\mathbf{r}')\mathrm{d}^3r'$$

where

$$V(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r}|\hat{V}|\mathbf{r}'\rangle$$

In spherical symmetry, we have

$$\langle \mathbf{r}|\hat{V}|\varphi\rangle = \langle \mathbf{r}|R^{-1}\hat{V}R|\varphi\rangle = \langle \mathbf{r}|R^{\dagger}\hat{V}R|\varphi\rangle = \int \langle \mathbf{r}|R^{\dagger}\hat{V}R|\mathbf{r}'\rangle \, \langle \mathbf{r}'|\varphi\rangle \, \mathrm{d}^3r' =$$

$$= \int \langle R\mathbf{r}|\hat{V}|R\mathbf{r}'\rangle \, \langle \mathbf{r}'|\varphi\rangle \, \mathrm{d}^3r' = \int V(R\mathbf{r}, R\mathbf{r}')\varphi(\mathbf{r}')\mathrm{d}^3r'$$

where R is the rotation operator (it's unitary). We have thus derived $V(R\mathbf{r}, R\mathbf{r}') = V(\mathbf{r}, \mathbf{r}')$ true for any R, which means that the the kernel only depends on ρ, ρ' and $\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'$, where $\mathbf{r} = \rho\hat{\mathbf{r}}$ and $\mathbf{r}' = \rho'\hat{\mathbf{r}}'$. So we obtain using (5.12)

$$V(\mathbf{r}, \mathbf{r}') = V(\rho, \rho', \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') = \sum_{lm} Y_{lm}(\hat{\mathbf{r}})V_l(\rho, \rho')Y^*_{lm}(\hat{\mathbf{r}}')$$

where

$$V_l(\rho,\rho') = \frac{(2l+1)^2}{8\pi}\int_{-1}^{1} P_l(x)V_l(\rho,\rho',x)\mathrm{d}x$$

In Dirac notation:

$$V(\mathbf{r},\mathbf{r}') = \langle\mathbf{r}|\hat{V}|\mathbf{r}'\rangle = \langle\hat{\mathbf{r}}|\langle\rho|\hat{V}|\rho'\rangle|\hat{\mathbf{r}}'\rangle = \sum_{lml'm'}\langle\hat{\mathbf{r}}|lm\rangle\langle lm|\langle\rho|\hat{V}|\rho'\rangle|l'm'\rangle\langle l'm'|\hat{\mathbf{r}}'\rangle$$

From the above derivation we see that we must have:

$$\langle lm|\langle\rho|\hat{V}|\rho'\rangle|l'm'\rangle = V_l(\rho,\rho')\delta_{ll'}\delta_{mm'}$$

in other words

$$V_l(\rho,\rho') = \langle lm|\langle\rho|\hat{V}|\rho'\rangle|lm\rangle \tag{3.32}$$

so we get

$$\langle\mathbf{r}|\hat{V}|\mathbf{r}'\rangle = \sum_{lm}\langle\hat{\mathbf{r}}|lm\rangle V_l(\rho,\rho')\langle lm|\hat{\mathbf{r}}'\rangle = \sum_{lm} Y_{lm}(\theta,\phi)V_l(\rho,\rho')Y^*_{lm}(\theta',\phi')$$

and

$$\hat{V} = \sum_{lm}|lm\rangle\langle lm|\hat{V}|lm\rangle\langle lm| = \sum_{lm}|lm\rangle\hat{V}_l\langle lm|$$

where the operator $\hat{V}_l = \langle lm|\hat{V}|lm\rangle$ only acts on the radial part of the wavefunction and according to (3.32) it doesn't depend on m. Also according to (3.32) its matrix elements are

$$\langle\rho|\hat{V}_l|\rho'\rangle = \langle lm|\langle\rho|\hat{V}|\rho'\rangle|lm\rangle = V_l(\rho,\rho')$$

3.3.3 Nonlocal Pseudopotentials

A nonlocal pseudopotential $\hat{V}$ is just a general Hermitian operator. We only want to construct pseudopotentials in the spherical problem, so every pseudopotential can be written in the form (3.31). In practice we only use either *local* (the operator $\hat{V}$ is local) or *semilocal* (the operator $\hat{V}$ is radially local, but angularly nonlocal) pseudopotential.

Local potential (radially and angularly local) is defined by:

$$\langle\mathbf{r}|\hat{V}|\mathbf{r}'\rangle = V(\rho)\langle\mathbf{r}|\mathbf{r}'\rangle$$

so we can simply write

$$\hat{V} = V(\rho) \tag{3.33}$$

so

$$V_l(\rho,\rho') = \langle lm|\,\langle\rho|\hat{V}|\rho'\rangle\,|lm\rangle = V(\rho)\,\langle\rho|\rho'\rangle = V(\rho)\frac{\delta(\rho-\rho')}{\rho^2}$$

so it turned out that the kernel is local and doesn't depend on l and we get

$$V(\mathbf{r}(\rho,\theta,\phi),\mathbf{r}'(\rho',\theta',\phi')) = \sum_{lm} Y_{lm}(\theta,\phi)V(\rho)\frac{\delta(\rho-\rho')}{\rho^2}Y^*_{lm}(\theta',\phi') =$$

$$= V(\rho)\frac{1}{\rho^2\sin\theta}\delta(\rho-\rho')\delta(\theta-\theta')\delta(\phi-\phi') = V(\rho)\delta(\mathbf{r}-\mathbf{r}')$$

and

$$\langle\mathbf{r}|\hat{V}|\varphi\rangle = \int V(\rho)\delta(\mathbf{r}-\mathbf{r}')\varphi(\mathbf{r}')\mathrm{d}^3r' = V(\rho)\varphi(\mathbf{r})$$

so we recover (3.33). There is no new information in these formulas, but it is useful to understand how to work with them and how to rewrite one to the other.

For a semilocal potential (radially local, but angularly nonlocal), the kernel cannot depend on m and is radially local, so:

$$\langle\rho|\hat{V}_l|\rho'\rangle = V_l(\rho,\rho') = \langle lm|\,\langle\rho|\hat{V}|\rho'\rangle\,|lm\rangle = V_l(\rho)\,\langle\rho|\rho'\rangle = V_l(\rho)\frac{\delta(\rho-\rho')}{\rho^2}$$

so the kernel is local and does depend on l and we simply write

$$\hat{V}_l = V_l(\rho)$$

and

$$\hat{V} = \sum_{lm} |lm\rangle\, V_l(\rho)\,\langle lm| \tag{3.34}$$

We can also calculate the same result explicitly in the $\mathbf{r}$ representation:

$$V(\mathbf{r}(\rho,\theta,\phi),\mathbf{r}'(\rho',\theta',\phi')) = \sum_{lm} Y_{lm}(\theta,\phi)V_l(\rho)\frac{\delta(\rho-\rho')}{\rho^2}Y^*_{lm}(\theta',\phi')$$

and

$$\langle\mathbf{r}|\hat{V}|\varphi\rangle = \int \sum_{lm} Y_{lm}(\theta,\phi)V_l(\rho)\frac{\delta(\rho-\rho')}{\rho^2}Y^*_{lm}(\theta',\phi')\varphi(\mathbf{r}')\mathrm{d}^3r' =$$

$$= \sum_{lm} Y_{lm}(\theta,\phi)V_l(\rho)\int Y^*_{lm}(\theta',\phi')\varphi(\rho\hat{\mathbf{r}}')\mathrm{d}\Omega'$$

or in Dirac notation

$$\langle \mathbf{r}|\hat{V}|\varphi\rangle = \sum_{lm} \langle \hat{\mathbf{r}}|lm\rangle \, V_l(\rho) \, \langle lm| \, \langle \rho|\varphi\rangle$$

and we recover (3.34).

So, to sum it up: semilocal pseudopotential is a general hermitian operator in the spherically symmetric problem (i.e. $\hat{V} = R^{-1}\hat{V}R$) and radially local. All such operators can be written in the form (3.34).

Now, it can be shown that if we make the assumption of radial locality, we get "correct" wavefunctions and energies in the linear approximation. We generally only take a few terms in the expansion (3.34), usually only V_0, V_1 and V_2, sometimes also V_3 and V_4.

3.3.4 Separable Potentials

The pseudopotential above (Hamman, Schlüter, Chiang) has the form

$$\hat{V} = \sum_{lm} |lm\rangle \, V_l(\rho) \, \langle lm| = V_{loc}(\rho) + \sum_{lm} |lm\rangle \, [V_l(\rho) - V_{loc}(\rho)] \, \langle lm|$$

Or, equivalently, in the $\mathbf{r}$ representation:

$$V(\mathbf{r},\mathbf{r}') = \langle \mathbf{r}|\hat{V}|\mathbf{r}'\rangle = V_{loc}(\rho)\delta(\mathbf{r}-\mathbf{r}') + \frac{\delta(\rho-\rho')}{\rho^2} \sum_{lm} Y_{lm}(\hat{\mathbf{r}})[V_l(\rho)-V_{loc}(\rho)]Y^*_{lm}(\hat{\mathbf{r}}')$$

The first term doesn't cause a problem. Let's denote the second term (which is semilocal) simply by v:

$$v = \sum_{lm} |lm\rangle \, [V_l(\rho) - V_{loc}(\rho)] \, \langle lm|$$

Let's choose a complete but otherwise arbitrary set of functions $|\phi_i\rangle$ (they contain both a radial and an angular dependence) and define a matrix U is by the equation

$$\sum_j U_{ij} \, \langle \phi_j|v|\phi_k\rangle = \delta_{ik}$$

then ($|\psi\rangle = |\phi_k\rangle \, \alpha_k$):

$$v \, |\psi\rangle = \sum_{ik} v \, |\phi_i\rangle \, \delta_{ik}\alpha_k = \sum_{ijk} v \, |\phi_i\rangle \, U_{ij} \, \langle \phi_j|v|\phi_k\rangle \, \alpha_k = \sum_{ij} v \, |\phi_i\rangle \, U_{ij} \, \langle \phi_j|v|\psi\rangle$$

So any Hermitian operator (including v) can be transformed exactly into the following form

$$v = \sum_{ij} v\,|\phi_i\rangle\, U_{ij}\, \langle\phi_j|\, v$$

We diagonalize the matrix U by choosing such functions $|\bar{\phi}_i\rangle$ for which the matrix $\langle\bar{\phi}_j|v|\bar{\phi}_k\rangle$ (and hence the corresponding matrix U) is equal to $\mathbb{1}$. We can find such functions for example using the Gram-Schmidt orthogonalization procedure on $|\phi_i\rangle$ with a norm $\langle f|v|g\rangle$ (for functions f and g), more on that later. Then

$$v = \sum_i v\,|\bar{\phi}_i\rangle\, \frac{1}{\langle\bar{\phi}_i|v|\bar{\phi}_i\rangle}\, \langle\bar{\phi}_i|\, v = \sum_i v\,|\bar{\phi}_i\rangle\, \langle\bar{\phi}_i|\, v \tag{3.35}$$

We could take any $|\phi_i\rangle$ and orthogonalize them. But because we have v in the form of (3.34), we will be using $|\phi_i\rangle$ in the form $|\phi_i\rangle = |R_{nl}\rangle\,|lm\rangle$, because it turns out we will only need to orthogonalize the radial parts. The first term in (3.35) then corresponds to the Kleinman-Bylander potential[11]. We of course take more terms and get accurate results without ghost states.

Let's look at the orthogonalization. We start with the wavefunctions:

$$|\phi_i\rangle = |R_{nl}\rangle\,|lm\rangle$$

where $R_{nl}(\rho) = \langle\rho|R_{nl}\rangle$ and i goes over all possible triplets (nlm).

We can also relate the i and n, l, m using this formula

$$i_{nlm} = \sum_{k=1}^{n-1} k^2 + \left(\sum_{k=0}^{l-1}(2k+1)\right) + (l+m+1) = \frac{(n-1)n(2n-1)}{6} + l(l+1) + m + 1$$

The operator v acts on these $|\phi_i\rangle$ like this

$$\langle\mathbf{r}|v|\phi_i\rangle = \langle\mathbf{r}|v|R_{nl}\rangle\,|lm\rangle = \langle\hat{\mathbf{r}}|\,\langle\rho|V_l(\rho)|R_{nl}\rangle\,|lm\rangle = V_l(\rho)R_{nl}(\rho)Y_{lm}(\hat{\mathbf{r}})$$

Now we need to construct new orthogonal set of functions $|\bar{\phi}_i\rangle$ satisfying

$$\langle\bar{\phi}_i|v|\bar{\phi}_j\rangle = \delta_{ij}$$

This can be done using several methods, we chose the Gram-Schmidt or-

thogonalization procedure, which works according to the following scheme:

$$
\begin{aligned}
|\tilde{\phi}_1\rangle &= \mathbb{1}\frac{1}{\sqrt{\langle\phi_1|v|\phi_1\rangle}}\,|\phi_1\rangle\,; & |\bar{\phi}_1\rangle &= \frac{1}{\sqrt{\langle\tilde{\phi}_1|v|\tilde{\phi}_1\rangle}}\,|\tilde{\phi}_1\rangle \\
|\tilde{\phi}_2\rangle &= \left(\mathbb{1} - |\bar{\phi}_1\rangle\,\langle\bar{\phi}_1|\,v\right)\frac{1}{\sqrt{\langle\phi_2|v|\phi_2\rangle}}\,|\phi_2\rangle\,; & |\bar{\phi}_2\rangle &= \frac{1}{\sqrt{\langle\tilde{\phi}_2|v|\tilde{\phi}_2\rangle}}\,|\tilde{\phi}_2\rangle \\
|\tilde{\phi}_3\rangle &= \left(\mathbb{1} - |\bar{\phi}_1\rangle\,\langle\bar{\phi}_1|\,v - |\bar{\phi}_2\rangle\,\langle\bar{\phi}_2|\,v\right)\frac{1}{\sqrt{\langle\phi_3|v|\phi_3\rangle}}\,|\phi_3\rangle\,; & |\bar{\phi}_3\rangle &= \frac{1}{\sqrt{\langle\tilde{\phi}_3|v|\tilde{\phi}_3\rangle}}\,|\tilde{\phi}_3\rangle \\
&\ldots
\end{aligned}
$$

We can verify by a direct calculation that this procedure ensures

$$\langle\bar{\phi}_i|v|\bar{\phi}_j\rangle = \delta_{ij}$$

It may be useful to compute the normalization factors explicitly:

$$
\begin{aligned}
\langle\tilde{\phi}_1|v|\tilde{\phi}_1\rangle &= 1 \\
\langle\tilde{\phi}_2|v|\tilde{\phi}_2\rangle &= 1 - \frac{\langle\phi_2|v|\bar{\phi}_1\rangle\,\langle\bar{\phi}_1|v|\phi_2\rangle}{\langle\phi_2|v|\phi_2\rangle} \\
\langle\tilde{\phi}_3|v|\tilde{\phi}_3\rangle &= 1 - \frac{\langle\phi_3|v|\bar{\phi}_1\rangle\,\langle\bar{\phi}_1|v|\phi_3\rangle + \langle\phi_3|v|\bar{\phi}_2\rangle\,\langle\bar{\phi}_2|v|\phi_3\rangle}{\langle\phi_3|v|\phi_3\rangle} \\
&\ldots
\end{aligned}
$$

we can also write down a first few orthogonal vectors explicitly:

$$
\begin{aligned}
|\bar{\phi}_1\rangle &= \frac{|\phi_1\rangle}{\sqrt{\langle\phi_1|v|\phi_1\rangle}} \\
|\bar{\phi}_2\rangle &= \frac{|\phi_2\rangle\,\langle\phi_1|v|\phi_1\rangle - |\phi_1\rangle\,\langle\phi_1|v|\phi_2\rangle}{\sqrt{\left(\langle\phi_1|v|\phi_1\rangle\,\langle\phi_2|v|\phi_2\rangle - \langle\phi_2|v|\phi_1\rangle\,\langle\phi_1|v|\phi_2\rangle\right)\langle\phi_1|v|\phi_1\rangle\,\langle\phi_2|v|\phi_2\rangle}}
\end{aligned}
$$

Now the crucial observation is

$$\langle lm|\,\langle R_{nl}|v|R_{n'l'}\rangle\,|l'm'\rangle = \langle R_{nl}|V_l(\rho)|R_{n'l}\rangle\,\delta_{ll'}\delta_{mm'}$$

which means that $\langle\phi_i|v|\phi_j\rangle = 0$ if $|\phi_i\rangle$ and $|\phi_j\rangle$ have different l or m. In other words $|\phi_i\rangle$ and $|\phi_j\rangle$ for different $|lm\rangle$ are already orthogonal. Thus

the G-S orthogonalization procedure only makes the R_{nl} orthogonal for the same $|lm\rangle$. To get explicit expressions for $|\bar{\phi}_i\rangle$, we simply use the formulas above and get:

$$|\phi_i\rangle = |R_{nl}\rangle\,|lm\rangle \quad \rightarrow \quad |\bar{\phi}_i\rangle = |\bar{R}_{nl}\rangle\,|lm\rangle$$

where we have constructed new $|\bar{R}_{nl}\rangle$ from original $|R_{nl}\rangle$:

$$\begin{aligned}
|\bar{R}_{10}\rangle &= \frac{|R_{10}\rangle}{\sqrt{\langle R_{10}|V_0|R_{10}\rangle}} \\
|\bar{R}_{20}\rangle &= \frac{|R_{20}\rangle - |\bar{R}_{10}\rangle\,\langle\bar{R}_{10}|V_0|R_{20}\rangle}{\sqrt{\cdots}} \\
|\bar{R}_{21}\rangle &= \frac{|R_{21}\rangle}{\sqrt{\langle R_{21}|V_1|R_{21}\rangle}} \\
|\bar{R}_{30}\rangle &= \frac{|R_{30}\rangle - |\bar{R}_{10}\rangle\,\langle\bar{R}_{10}|V_0|R_{30}\rangle - |\bar{R}_{20}\rangle\,\langle\bar{R}_{20}|V_0|R_{30}\rangle}{\sqrt{\cdots}} \\
|\bar{R}_{31}\rangle &= \frac{|R_{31}\rangle - |\bar{R}_{21}\rangle\,\langle\bar{R}_{21}|V_1|R_{31}\rangle}{\sqrt{\cdots}} \\
|\bar{R}_{32}\rangle &= \frac{|R_{32}\rangle}{\sqrt{\langle R_{32}|V_1|R_{32}\rangle}} \\
|\bar{R}_{40}\rangle &= \frac{|R_{40}\rangle - |\bar{R}_{10}\rangle\,\langle\bar{R}_{10}|V_0|R_{40}\rangle - |\bar{R}_{20}\rangle\,\langle\bar{R}_{20}|V_0|R_{40}\rangle - |\bar{R}_{30}\rangle\,\langle\bar{R}_{30}|V_0|R_{40}\rangle}{\sqrt{\cdots}} \\
|\bar{R}_{41}\rangle &= \frac{|R_{41}\rangle - |\bar{R}_{21}\rangle\,\langle\bar{R}_{21}|V_1|R_{41}\rangle - |\bar{R}_{31}\rangle\,\langle\bar{R}_{31}|V_1|R_{41}\rangle}{\sqrt{\cdots}} \\
&\cdots
\end{aligned}$$

Ok, so we have constructed new $|\bar{R}_{nl}\rangle$ from $|R_{nl}\rangle$ which obey

$$\langle\bar{R}_{nl}|V_l|\bar{R}_{n'l}\rangle = \delta_{nn'} \tag{3.36}$$

so for every V_l, we construct $|\bar{R}_{nl}\rangle$ for $n = l+1,\ l+2, \cdots$. Let's continue:

$$v\,|\bar{\phi}_i\rangle = V_l(\rho)\,|\bar{R}_{nl}\rangle\,|lm\rangle$$

and finally we arrive at the separable form of the l dependent pseudopotential

$$v = \sum_i v\,|\bar{\phi}_i\rangle\,\langle\bar{\phi}_i|\,v = \sum_i V_l(\rho)\,|\bar{R}_{nl}\rangle\,|lm\rangle\,\langle lm|\,\langle\bar{R}_{nl}|\,V_l(\rho) \tag{3.37}$$

Note: the V_l is actually $V_l - V_{loc}$, but this is just a detail.

To have some explicit formula, let's write how the separable potential acts on a wavefunction:

$$(v\psi)(\mathbf{r}) = \langle \mathbf{r}|v|\psi\rangle = \sum_i \langle \hat{\mathbf{r}}|lm\rangle \langle \rho|V_l(\rho)|\bar{R}_{nl}\rangle \langle \bar{R}_{nl}| V_l(\rho) \langle lm|\psi\rangle =$$

$$= \sum_i Y_{lm}(\hat{\mathbf{r}})\bar{R}_{nl}(\rho)V_l(\rho) \int \bar{R}_{nl}(\rho')V_l(\rho') \int Y^*_{lm}(\hat{\mathbf{r}}')\psi(\mathbf{r}')\,\mathrm{d}\Omega'\,\rho'^2\mathrm{d}\rho' =$$

$$= \sum_i Y_{lm}(\hat{\mathbf{r}})\bar{R}_{nl}(\rho)V_l(\rho) \int \bar{R}_{nl}(\rho')V_l(\rho')Y^*_{lm}(\hat{\mathbf{r}}')\psi(\mathbf{r}')\,\mathrm{d}^3r'$$

To have some insight on what we are actually doing: we are making the local potential V_l nonlocal using:

$$V_l = \sum_{n=l+1}^{\infty} V_l\,|\bar{R}_{nl}\rangle \langle \bar{R}_{nl}|\,V_l \tag{3.38}$$

where

$$\langle \bar{R}_{nl}|V_l|\bar{R}_{n'l}\rangle = \delta_{nn'}$$

or in $\mathbf{r}$ representation:

$$V_l(\rho)\psi(\rho\hat{\mathbf{r}}) = \sum_n V_l(\rho)\bar{R}_{nl}(\rho) \int \bar{R}_{nl}(\rho')V_l(\rho')\psi(\rho'\hat{\mathbf{r}})\rho'^2\mathrm{d}\rho'$$

which is useful when computing integrals of this type

$$V_{ij} = \int \phi_i(\rho)V_l\phi_j(\rho)\rho^2\mathrm{d}^3\rho = \langle i|V_l|j\rangle = \sum_n \langle i|V_l|\bar{R}_{nl}\rangle \langle \bar{R}_{nl}|V_l|j\rangle$$

$$\langle i|V_l|\bar{R}_{nl}\rangle = \int \phi_i(\rho)V_l(\rho)\bar{R}_{nl}(\rho)\rho^2\mathrm{d}\rho$$

because the integral on the left hand side actually represents N^2 integrals, where N is the number of basis vectors $|\phi_i\rangle$. The sum on the right hand side however only represents $K \cdot N$ integrals, where K is the number of terms taken into account in (3.38). Of course taking only finite number of terms in (3.38) is only an approximation to $\hat{V}_l$. In our case, we don't need these 1D integrals (which can be easily computed directly, because V_l is local and the basis functions ϕ_i are nonzero only around a node in the mesh, which

means that the matrix V_{ij} is sparse), but 3D integrals, where angular parts of V are nonlocal and radial part is local (so the matrix V_{ij} is dense), so the above procedure is the only way how to proceed, because we decompose the matrix V_{ij} into the sum of matrices in the form $p_i p_j^*$, which can easily be handled and solved.

The scheme for the separation described above works for any functions $R_{nl}(\rho)$. Because of the form of the expansion (3.38) however, we will use R_{nl} from one atomic calculation. We need to approximate V_l by as few terms as possible, so imagine how the $V_l(\rho)$ acts on the lowest radial function in the l subspace, which is $|R_{l+1;l}\rangle$ and we see that all the terms in (3.38) except the first one $V_l \left|\bar{R}_{l+1;l}\right\rangle \left\langle\bar{R}_{l+1;l}\right| V_l$ give zero, because they are orthogonal to $|R_{l+1;l}\rangle$. For the function $|R_{l+2;l}\rangle$ all the terms except the first two are zero, because $\langle\bar{R}_{nl}|V_0|R_{l+2;l}\rangle \neq 0$ only for $n = l+1$ or $n = l+2$ (because the vectors $|R_{l+1;l}\rangle$ and $|R_{l+2;l}\rangle$ span the same subspace as $|\bar{R}_{l+1;l}\rangle$ and $|\bar{R}_{l+2;l}\rangle$ and using (3.36)) For functions, which are a little different from all $|R_{nl}\rangle$ ($n > l$), we won't generally get precise results taking any (finite) number of terms in (3.38), but the higher terms should give smaller and smaller corrections.

So, to sum it up: We take all the V_l in (3.37) as we did in (3.34). Theoretically we should take $\bar{R}_{nl}$ for all $n = l+1,\ l+2,\ l+3, \ldots$, but practically it suffices to only take several $\bar{R}_{nl}$ for a given l from one atomic calculation.

Let's give an example: we are calculating 14 electrons, so we will only take into account the lowest 14 eigenvalues in the Kohn sham equations, which are $|\phi_1\rangle$ up to $|\phi_{14}\rangle$. The lowest radial functions in each l subspace are $|\phi_i\rangle$ for $i = 1, 3, 4, 5, 10, 11, 12, 13, 14$ and on these 9 functions we get a precise result with only one term in the expansion (3.38). For the other 5 functions ($i = 2, 6, 7, 8, 9$) we will have to take into account more terms. Let's look in more detail at the case $l = 0$ (i.e. $i = 1, 2, 6$). Then

$$V_0 = V_0 \left|\bar{R}_{10}\right\rangle \left\langle\bar{R}_{10}\right| V_0 + V_0 \left|\bar{R}_{20}\right\rangle \left\langle\bar{R}_{20}\right| V_0 + V_0 \left|\bar{R}_{30}\right\rangle \left\langle\bar{R}_{30}\right| V_0 + \ldots$$

and for the case $i = 1$ we see that one term in (3.38) is enough:

$$v\left|\phi_1\right\rangle = v\left|R_{10}\right\rangle\left|00\right\rangle = V_0\left|R_{10}\right\rangle\left|00\right\rangle = V_0\left|\bar{R}_{10}\right\rangle\left\langle\bar{R}_{10}\right| V_0\left|R_{10}\right\rangle\left|00\right\rangle$$

because $\langle\bar{R}_{n0}|V_0|R_{10}\rangle = 0$ for $n > 1$. For the case $i = 2$ we get the correct result with 2 terms in (3.38)

$$v\left|\phi_2\right\rangle = v\left|R_{20}\right\rangle\left|00\right\rangle = V_0\left|R_{20}\right\rangle\left|00\right\rangle = \left(V_0\left|\bar{R}_{10}\right\rangle\left\langle\bar{R}_{10}\right| V_0\left|R_{20}\right\rangle + V_0\left|\bar{R}_{20}\right\rangle\left\langle\bar{R}_{20}\right| V_0\left|R_{20}\right\rangle\right)\left|00\right\rangle$$

because $\langle \bar{R}_{n0}|V_0|R_{20}\rangle = 0$ for $n > 2$. For the case $i = 6$ we need to take into account 3 terms etc. We can see from this example, that taking $|R_{nl}\rangle$ from one atomic calculation, we get precise results (with the same atom) only taking into account a finite number of terms in (3.38), for 14 electrons actually only 3 terms. For several atoms calculation, we won't get precise results, but it should be a sufficiently good approximation.

The described method is general, the only drawback is that if we don't take functions $|R_{nl}\rangle$ which are similar to the solution, we need to take a lot of terms in (3.38), resulting in many matrices of the form $p_i p_j^*$, which we don't want, even though, theoretically we can get a solution with any precision we want taking more and more terms in (3.38).

See also [2].

3.4 Finite Element Method

The finite element method (FEM) is the most widely used technique for solving partial differential equations (PDE). In this section we explain how it can be applied to the Schrödinger equation.

3.4.1 Weak Formulation of the Schrödinger Equation

One-particle Schrödinger equation has the form

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi = E\psi \quad \text{in } \Omega, \tag{3.39}$$

where Ω is a Cartesian square (in 2D) or cube (in 3D) centered at the origin of the coordinate system. The domain is chosen sufficiently large so that the decay of ψ allows us to prescribe zero Dirichlet conditions on the boundary Γ. Here V is a given function in Ω representing the potential well.

We multiply both sides of (3.39) by a test function v that is zero on Γ and integrate over Ω,

$$\int_\Omega -\left(\frac{\hbar^2}{2m}\nabla^2\psi\right) v \,\mathrm{d}V = \int_\Omega (E - V)\psi v \,\mathrm{d}V\,. \tag{3.40}$$

Applying to the left-hand side the Green's first identity (generalized integration by parts formula), we obtain

$$\int_\Omega \frac{\hbar^2}{2m}\nabla\psi\cdot\nabla v \,\mathrm{d}V = \int_\Omega (E - V)\psi v \,\mathrm{d}V + \oint_\Gamma \frac{\hbar^2}{2m}(\nabla\psi)v\cdot\mathbf{n}\,\mathrm{d}S\,.$$

Since $v \equiv 0$ on Γ, the integral over Γ vanishes and we obtain the following weak formulation of (3.39): Find a function $\psi \in S$ such that

$$\int_\Omega \frac{\hbar^2}{2m} \nabla\psi \cdot \nabla v \, \mathrm{d}V = \int_\Omega (E - V)\psi v \, \mathrm{d}V \quad \text{for every } v \in S. \tag{3.41}$$

Here, S is a subspace of the standard Sobolev space $H^1(\Omega)$ corresponding to zero Dirichlet boundary conditions on Γ, i.e.,

$$S = \{v \in H^1(\Omega);\ v = 0 \text{ on } \Gamma\}.$$

Recall that the function space $H^1(\Omega)$ consists of all functions defined in Ω which are square-integrable and have square-integrable (weak) gradients.

3.4.2 Finite Element Method

The domain Ω is covered with a finite element mesh consisting of M non-overlapping triangles (in 2D) or tetrahedra (in 3D) $K_1, K_2, \ldots, K_M$. Any two elements can only share a common vertex, a common edge (along with the corresponding pair of vertices), or a common face (along with the edges and vertices corresponding to that face).

The finite element approximation takes place in a finite-dimensional subspace $S_n \subset S$ consisting of continuous functions defined in Ω which are linear within every element K_m, $m = 1, 2, \ldots, M$. Let ϕ_i, $i = 1, 2, \ldots, n$ be a basis of S_n. The approximate solution ψ is sought as a linear combination of these basis functions with unknown coefficients,

$$\psi = \sum q_j \phi_j.$$

Substituting this expansion into (3.41) and using the basis functions ϕ_i, $i = 1, 2, \ldots, n$ in place of the test function v, we obtain a discrete problem

$$\left(\int_\Omega \frac{\hbar^2}{2m} \nabla\phi_j \cdot \nabla\phi_i \, \mathrm{d}V + \int_\Omega \phi_i V \phi_j \, \mathrm{d}V \right) q_j = \left(\int_\Omega E \phi_j \phi_i \, \mathrm{d}V \right) q_j. \tag{3.42}$$

This can be written in a matrix form

$$(K_{ij} + V_{ij})\, q_j = E M_{ij} q_j \,,$$

where

$$\begin{aligned} V_{ij} &= \int_\Omega \phi_i V \phi_j \, \mathrm{d}V \,, \\ M_{ij} &= \int_\Omega \phi_i \phi_j \, \mathrm{d}V \,, \\ K_{ij} &= \frac{\hbar^2}{2m} \int_\Omega \nabla\phi_i \cdot \nabla\phi_j \, \mathrm{d}V \,. \end{aligned}$$

The integrals in (3.42) are evaluated via element contributions over $K_1, K_2, \ldots, K_M$. For example,

$$K_{ij} = \sum_{m=1}^{M} K_{ij}^{(m)},$$

where

$$K_{ij}^{(m)} = \int_{K_m} \frac{\hbar^2}{2m} \nabla\phi_j \cdot \nabla\phi_i \, \mathrm{d}V^E \approx \sum_{q=0}^{N_q - 1} \frac{\hbar^2}{2m} \nabla\phi_i(x_q) \cdot \nabla\phi_j(x_q) \, w_q |\det J(\hat{x}_q)| \,.$$

The integral is computed numerically using the Gauss integration technique: x_q are Gauss points (there are N_q of them), w_q is the weight of each point, and the Jacobian $|\det J(\hat{x}_q)|$ is there because we are actually computing the integral on the reference element instead in the real space. It is worth mentioning that for both triangular elements in 2D and tetrahedra in 3D, the Jacobian is constant.

Finite elements programs usually have an assembly phase, where they construct all the matrices, in our case:

$$(K_{ij} + V_{ij}) \, q_j = E M_{ij} q_j \,,$$

and then a solve phase. In our case this amounts to solving a large sparse generalized eigenvalue problem.

Chapter 4

Results

In this section we show results of our program. The program is opensource (BSD licensed) and available at `http://sfepy.org`, use the release 00.50.00 to get the results as below.

In most examples we present also an exact analytic solution (in the column denoted as "exact"). The particular way of obtaining the analytic results is described at each individual example.

4.1 Solution of the Schrödinger Equation

The first thing we need to do is to create a mesh.

4.1.1 2D: Mesh

We create a mesh:

```
$ ./schroedinger.py --mesh --2d
Dimension: 2
Info    : 'gmsh -2 tmp/mesh.geo -format mesh'
   started on Sat Jul 19 02:04:41 2008
Info    : Reading 'tmp/mesh.geo'
Info    : Read 'tmp/mesh.geo'
Info    : Meshing 1D...
Info    : Meshing curve 1 (Line)
Info    : Meshing curve 2 (Line)
Info    : Meshing curve 3 (Line)
Info    : Meshing curve 4 (Line)
```

```
Info    : Mesh 1D complete (0.028002 s)
Info    : Mesh
Info    : Meshing 2D...
Info    : Meshing surface 6 (Plane, MeshAdapt+Delaunay)
Info    : Mesh 2D complete (3.65223 s)
Info    : Mesh
Info    : 8100 vertices 16198 elements
Info    : Writing 'tmp/mesh.mesh'
Info    : Wrote 'tmp/mesh.mesh'
Mesh written to tmp/mesh.vtk
```

As you can see, we call gmsh[9] in the background to create a triangular mesh in 2D. See the fig. 4.1 and 4.2 for a visualization of the mesh in the paraview[10] program.

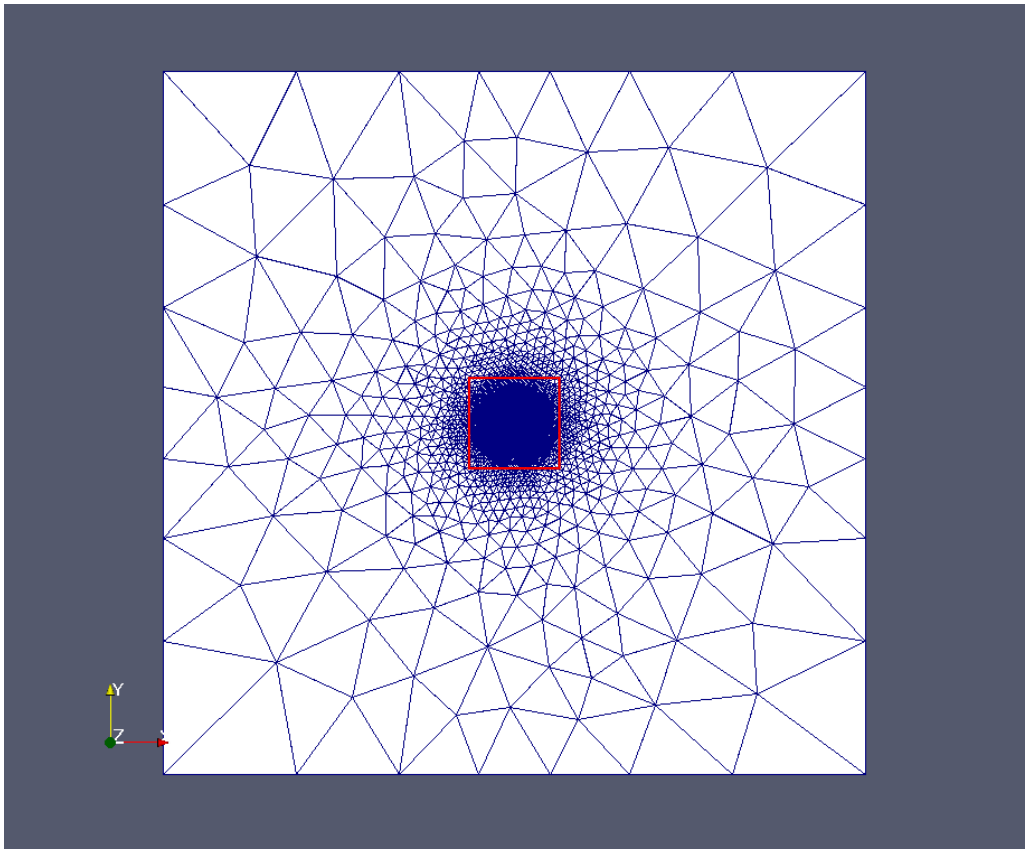


Figure 4.1: 2D mesh, the red square is zoomed in the fig. 4.2

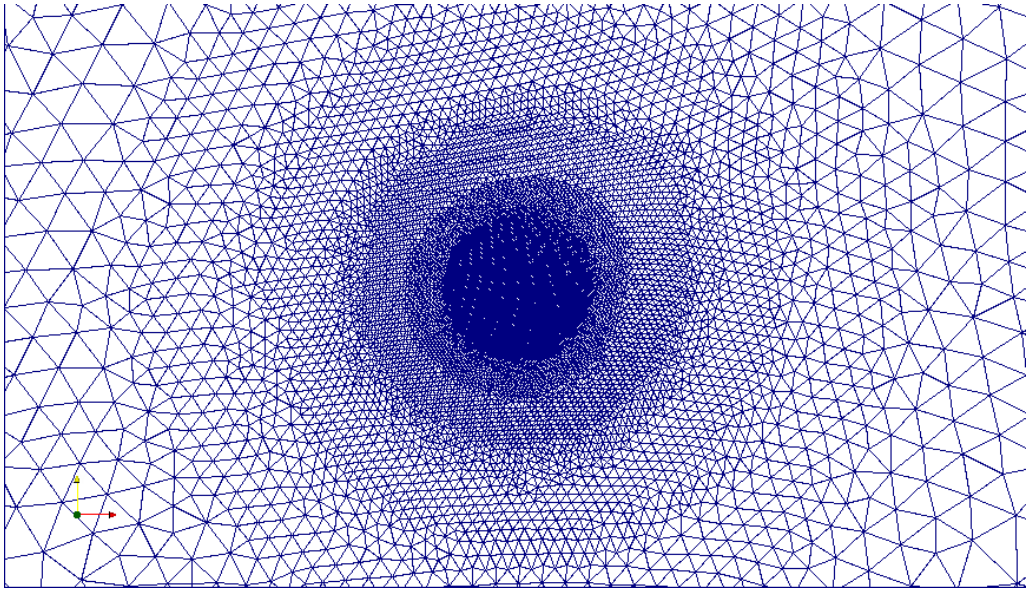


Figure 4.2: Zoomed in 2D mesh

4.1.2 2D: Potential Well

This is also sometimes called particle in the box.

$$V(x) = \begin{cases} 0, & \text{inside the square} \quad a \times a \\ \infty, & \text{outside} \end{cases}$$

Analytic solution:

$$E_{n_1 n_2} = \frac{\pi^2}{2a^2}\left(n_1^2 + n_2^2\right)$$

where $n_i = 1, 2, 3, \ldots$ are independent quantum numbers. For example for $a = 1$ we get: $E_{11} = 2$, $E_{12} = E_{21} = 5$, $E_{22} = 8$, $E_{13} = E_{31} = 10$, $E_{23} = E_{32} = 13$, $E_{14} = E_{41} = 17$, $E_{22} = 18$, $E_{24} = E_{42} = 20$, ...

Our program constructs the element matrices as described before and then calls PySparse[8] to solve the generalized eigenproblem.

```
$ ./schroedinger.py --well
Dimension: 2
[...]
Energies:
```

```
n       exact         FEM       error
0:   0.00098696   0.00100813    2.14%
1:   0.00246740   0.00255738    3.65%
2:   0.00246740   0.00256454    3.94%
3:   0.00394784   0.00421025    6.65%
4:   0.00493480   0.00524481    6.28%
5:   0.00493480   0.00525660    6.52%
6:   0.00641524   0.00705958   10.04%
7:   0.00641524   0.00706794   10.17%
8:   0.00838916   0.00916761    9.28%
9:   0.00838916   0.00920827    9.76%
Solution saved to mesh.vtk
```

The lowest 9 eigenvectors are visualized in fig. 4.3, from left to right, top to bottom.

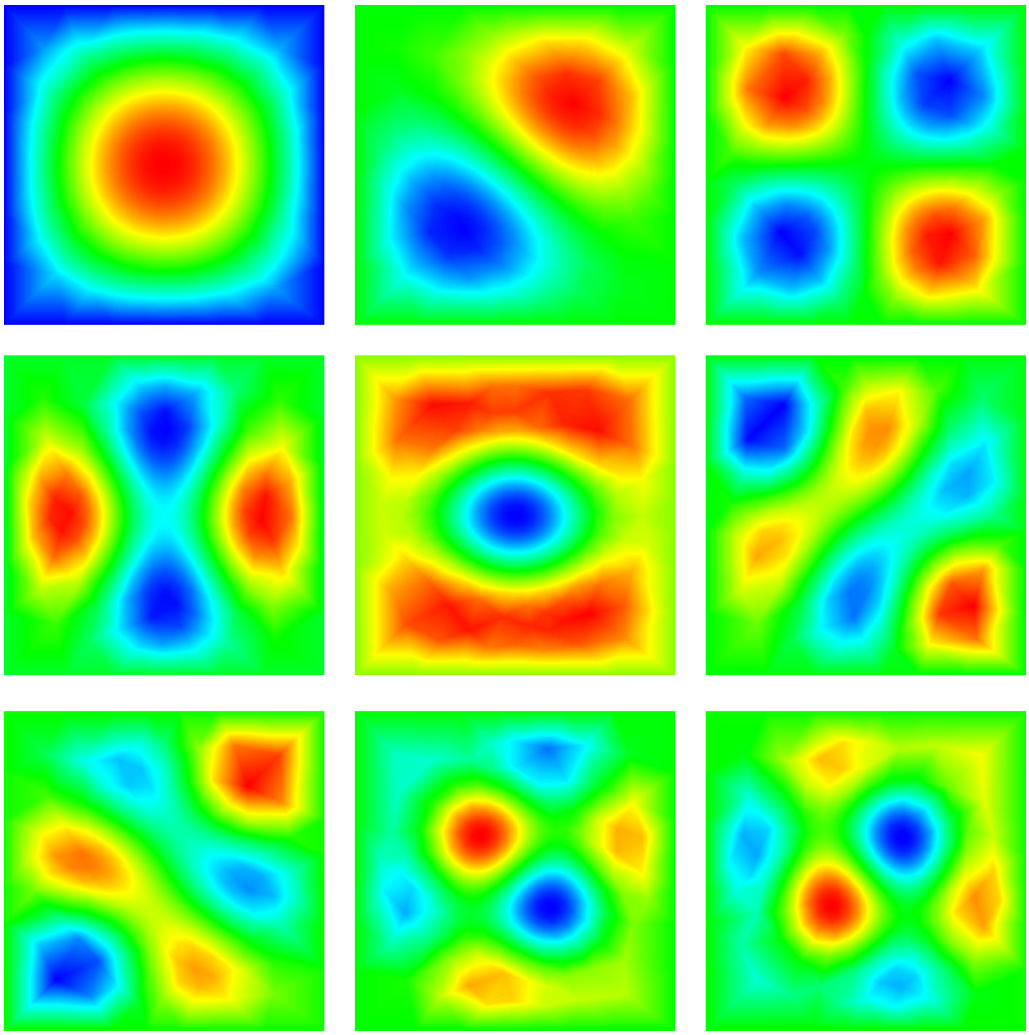


Figure 4.3: 2D well: eigenvectors 1 to 9

4.1.3 2D: Linear Harmonic Oscillator

$$V(r) = \begin{cases} \frac{1}{2}\omega^2 r^2, & \text{inside the square} \quad a \times a \\ \infty, & \text{outside} \end{cases}$$

Analytic solution in the limit $a \to \infty$:

$$E_{n_1 n_2} = (n_1 + n_2 + 1)\,\omega$$

where $n_1, n_2 = 0, 1, 2, \ldots$, so: $E_{00} = 1$, $E_{01} = E_{10} = 2$, $E_{11} = E_{20} = E_{02} = 3$, $E_{12} = E_{21} = E_{30} = E_{03} = 4$, $E_{22} = E_{13} = E_{31} = E_{40} = E_{04} = 5$,

```
$ ./schroedinger.py --oscillator
Dimension: 2
[...]
Energies:
n       exact          FEM          error
0:   1.00000000   1.00081703   0.08%
1:   2.00000000   2.00158339   0.08%
2:   2.00000000   2.00175346   0.09%
3:   3.00000000   3.00302269   0.10%
4:   3.00000000   3.00334897   0.11%
5:   3.00000000   3.00346025   0.12%
6:   4.00000000   4.00498898   0.12%
7:   4.00000000   4.00571512   0.14%
8:   4.00000000   4.00606336   0.15%
9:   4.00000000   4.00631703   0.16%
10:  5.00000000   5.00743227   0.15%
11:  5.00000000   5.00868813   0.17%
12:  5.00000000   5.00947456   0.19%
13:  5.00000000   5.01031228   0.21%
14:  5.00000000   5.01042976   0.21%
15:  6.00000000   6.01161255   0.19%
Solution saved to mesh.vtk
```

All the eigenvectors are visualized in fig. 4.4.

To get a better insight, we plotted the 8th and 10th eigenvector as a 2D surface in 3D, see the fig 4.5.

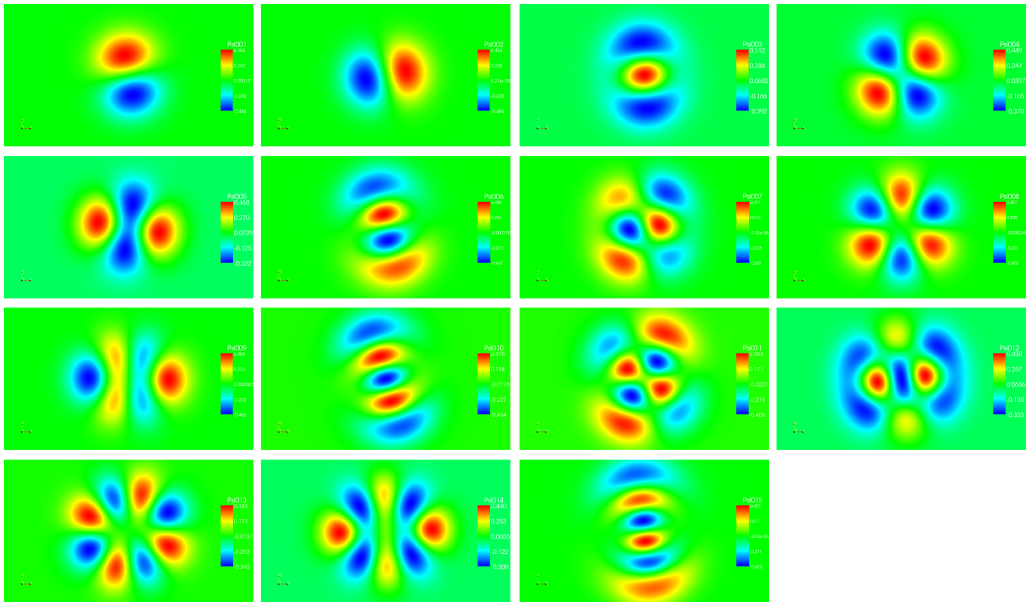


Figure 4.4: 2D oscillator: eigenvectors 1 to 15

4.1.4 2D: Hydrogen Atom

$$V(r) = \begin{cases} -\frac{Z^2}{r}, & \text{inside the square} \quad a \times a \\ \infty, & \text{outside} \end{cases}$$

Analytic solution in the limit $a \to \infty$:

$$E_n = -\frac{Z^2}{8(n-\frac{1}{2})^2}$$

where $n = 1, 2, 3, \ldots$. For $Z = 1$: $E_1 = -\frac{1}{2} = -0.5$, $E_2 = -\frac{1}{18} = -0.055$, $E_3 = -\frac{1}{50} = -0.020$. $E_4 = -\frac{1}{98} = -0.010$.

```
$ ./schroedinger.py --hydrogen
Dimension: 2
[...]
Energies:
n       exact          FEM          error
0:  -0.50000000  -0.48444312   3.11%
1:  -0.05555556  -0.05546247   0.17%
2:  -0.05555556  -0.05545794   0.18%
```

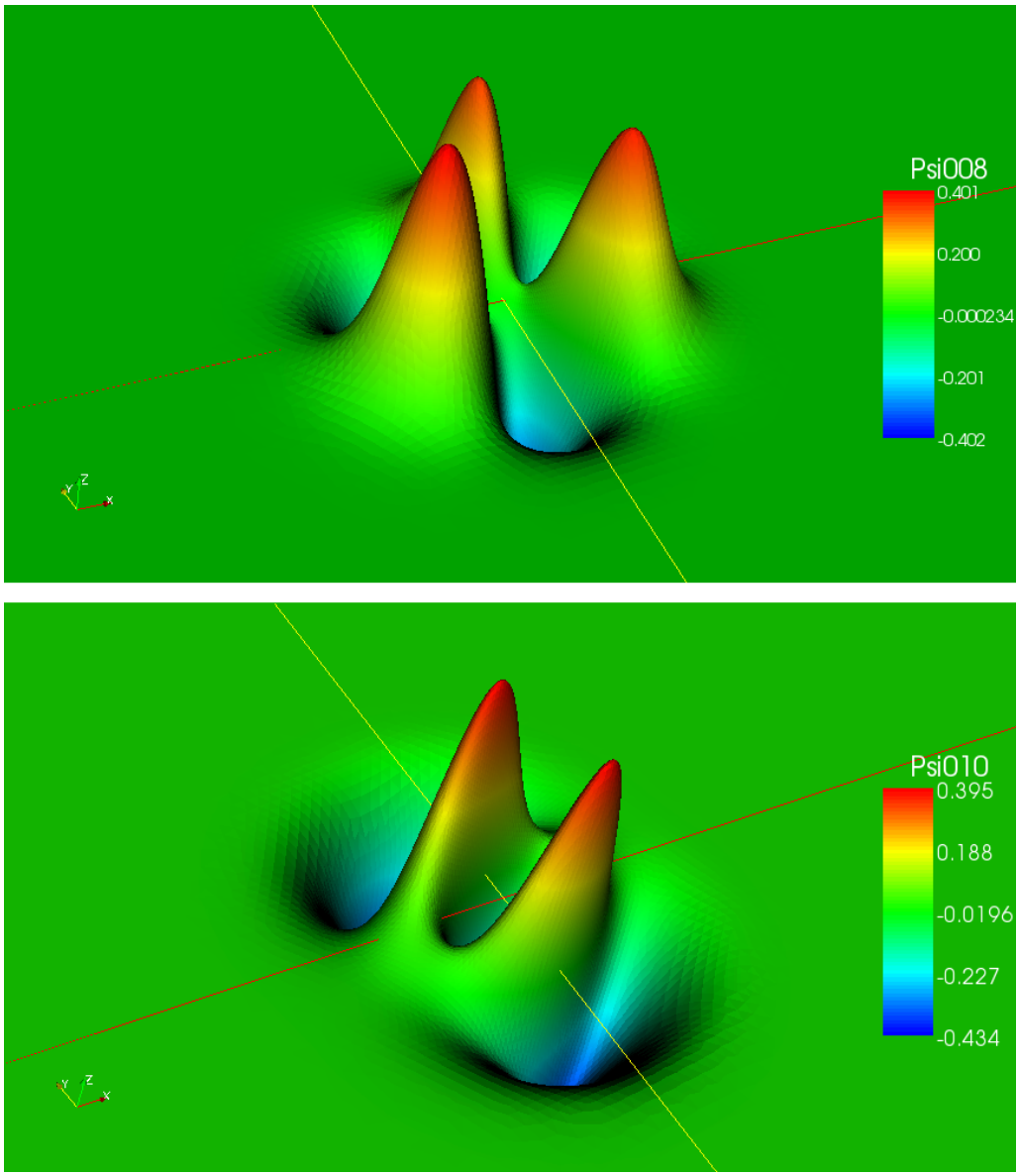


Figure 4.5: 2D oscillator: eigenvector 8 and 10, plotted as a surface

```
3:   -0.05555556    -0.05490833    1.17%
4:   -0.02000000    -0.01987759    0.61%
5:   -0.02000000    -0.01986552    0.67%
6:   -0.02000000    -0.01978261    1.09%
7:   -0.02000000    -0.01974034    1.30%
8:   -0.02000000    -0.01973349    1.33%
9:   -0.01020408    -0.00961296    5.79%
Solution saved to mesh.vtk
```

4.1.5 2D: Boron Atom

The analytic solution is the same as for the hydrogen atom, only with $Z = 5$, so each energy needs to by multiplied by 25.

```
$ ./schroedinger.py --boron
[...]
Energies:
n       exact           FEM       error
0:   -12.50000000   -12.65582408    1.25%
1:   -1.38888889    -1.39403240    0.37%
2:   -1.38888889    -1.38823512    0.05%
3:   -1.38888889    -1.38821053    0.05%
4:   -0.50000000    -0.50042416    0.08%
5:   -0.50000000    -0.49898923    0.20%
6:   -0.50000000    -0.49895079    0.21%
7:   -0.50000000    -0.49809201    0.38%
8:   -0.50000000    -0.49804906    0.39%
9:   -0.25510204    -0.25394635    0.45%
10:  -0.25510204    -0.25329618    0.71%
11:  -0.25510204    -0.25324683    0.73%
12:  -0.25510204    -0.25249085    1.02%
13:  -0.25510204    -0.25242286    1.05%
Solution saved to mesh.vtk
```

4.1.6 3D: Mesh

In 3D, we generate tetrahedra:

```
$ ./schroedinger.py --mesh
```

```
Dimension: 3
[...]
Mesh written to tmp/mesh.vtk
```

Look at the fig. 4.6 to see how the mesh looks like.

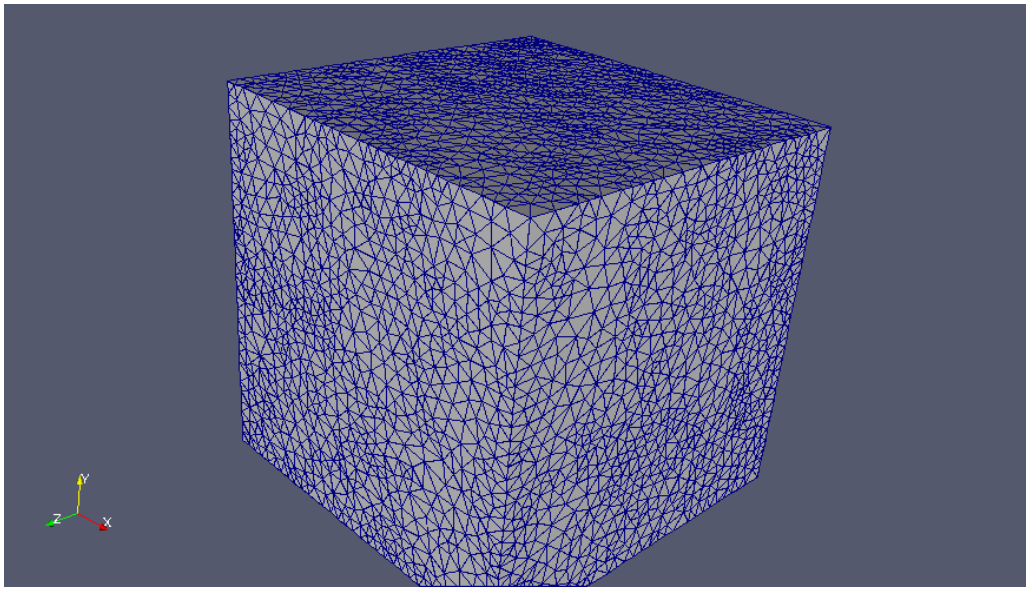


Figure 4.6: 3D mesh

4.1.7 3D: Potential Well

$$V(x) = \begin{cases} 0, & \text{inside the box} \quad a \times a \times a \\ \infty, & \text{outside} \end{cases}$$

Analytic solution:

$$E_{n_1 n_2 n_3} = \frac{\pi^2}{2a^2}\left(n_1^2 + n_2^2 + n_3^2\right)$$

where $n_i = 1, 2, 3, \ldots$ are independent quantum numbers. For example for $a = 1$ we get: $E_{111} = 14.804$, $E_{211} = E_{121} = E_{112} = 29.608$, $E_{122} = E_{212} = E_{221} = 44.413$, $E_{311} = E_{131} = E_{113} = 54.282$ $E_{222} = 59.217$, $E_{123} = E_{\text{perm.}} = 69.087$.

```
$ ./schroedinger.py --well
Dimension: 3
```

```
[...]
Energies:
n       exact         FEM          error
0:   0.14804407   0.14922535   0.80%
1:   0.29608813   0.30079010   1.59%
2:   0.29608813   0.30082698   1.60%
3:   0.29608813   0.30084093   1.61%
4:   0.44413220   0.45473187   2.39%
5:   0.44413220   0.45482735   2.41%
6:   0.44413220   0.45489304   2.42%
7:   0.54282824   0.55869467   2.92%
8:   0.54282824   0.55871268   2.93%
9:   0.54282824   0.55889291   2.96%
10:  0.59217626   0.61113461   3.20%
11:  0.69087231   0.71656267   3.72%
12:  0.69087231   0.71661781   3.73%
13:  0.69087231   0.71678728   3.75%
14:  0.69087231   0.71695827   3.78%
15:  0.69087231   0.71706618   3.79%
16:  0.69087231   0.71729282   3.82%
17:  0.83891637   0.87708704   4.55%
18:  0.83891637   0.87715932   4.56%
19:  0.83891637   0.87797932   4.66%
Solution saved to mesh.vtk
```

As you can see above, we got ($a = 1$, 24702 nodes):

E	1	2-4	5-7	8-10	11	12-
theory	14.804	29.608	44.413	54.282	59.217	69.087
FEM	14.861	29.833	44.919	55.035	60.123	70.305
		29.834	44.920	55.042		70.310
		29.836	44.925	55.047		...

So we got the correct energies and correct degeneracies. See the figs 4.7 – 4.9 for examples of eigenvectors.

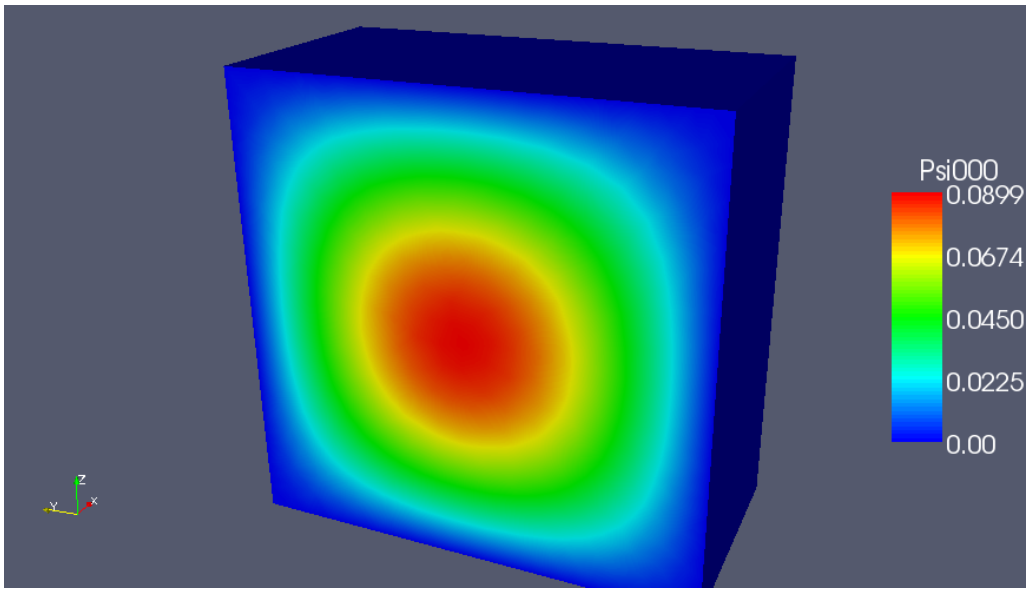


Figure 4.7: 3D potential well: eigenvector 1, cut plane

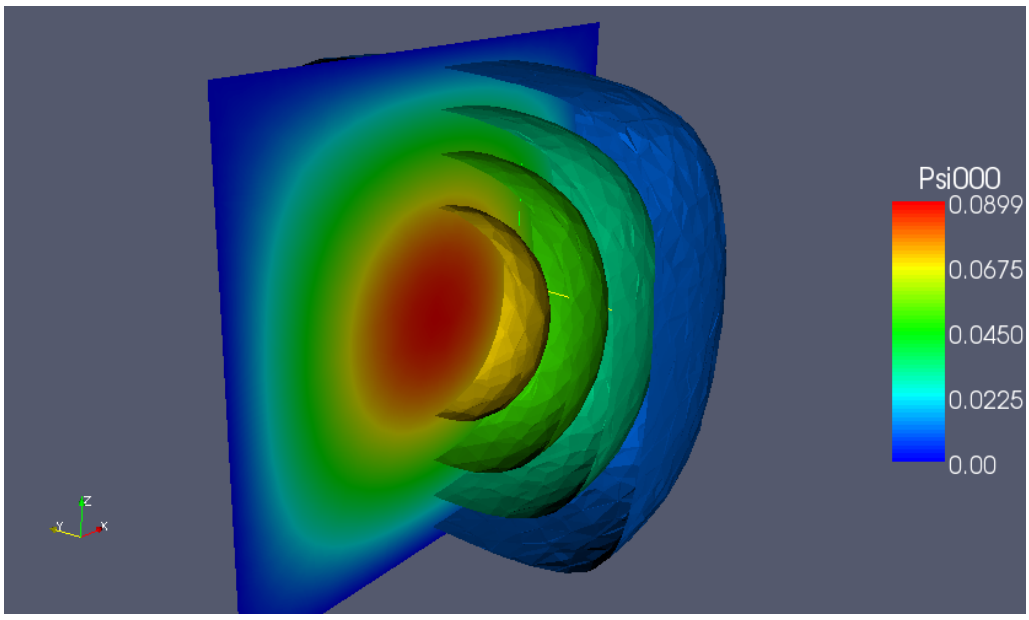


Figure 4.8: 3D potential well: eigenvector 1, cut plane and contours

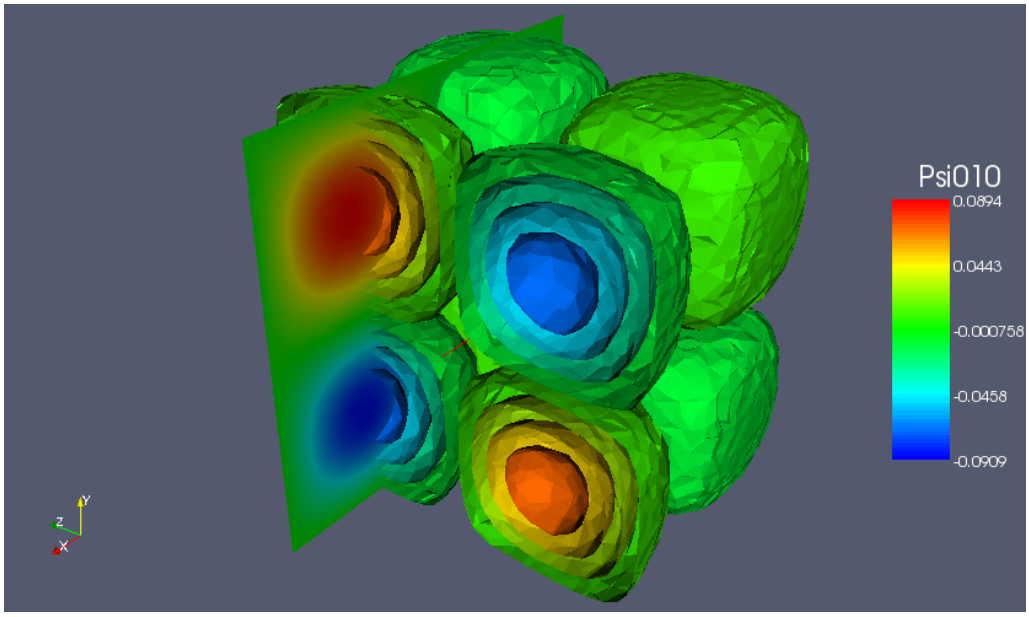


Figure 4.9: 3D potential well: eigenvector 10, cut plane and contours

4.1.8 3D: Linear Harmonic Oscillator

$$V(r) = \begin{cases} \frac{1}{2}\omega^2 r^2, & \text{inside the box} \quad a \times a \times a \\ \infty, & \text{outside} \end{cases}$$

Analytic solution in the limit $a \to \infty$:

$$E_{nl} = \left(2n + l + \frac{3}{2}\right)\omega$$

where $n, l = 0, 1, 2, \ldots$. Degeneracy is $2l + 1$, so: $E_{00} = \frac{3}{2}$, triple $E_{01} = \frac{5}{2}$, $E_{10} = \frac{7}{2}$, quintuple $E_{02} = \frac{7}{2}$, triple $E_{11} = \frac{9}{2}$, quintuple $E_{12} = \frac{11}{2}$:

```
$ ./schroedinger.py --oscillator
Dimension: 3
[...]
Energies:
n       exact          FEM          error
0:   1.50000000   1.60246586   6.83%
1:   2.50000000   2.66350384   6.54%
2:   2.50000000   2.66592179   6.64%
```

```
3:   2.50000000    2.66745129    6.70%
4:   3.50000000    3.73764474    6.79%
5:   3.50000000    3.74251455    6.93%
6:   3.50000000    3.74545592    7.01%
7:   3.50000000    3.74824003    7.09%
8:   3.50000000    3.75739243    7.35%
9:   3.50000000    3.78188467    8.05%
10:   4.50000000    4.84098311    7.58%
11:   4.50000000    4.84411120    7.65%
12:   4.50000000    4.84695596    7.71%
13:   4.50000000    4.85068992    7.79%
14:   4.50000000    4.85439463    7.88%
15:   4.50000000    4.86152828    8.03%
16:   4.50000000    4.86513080    8.11%
17:   4.50000000    4.91355134    9.19%
18:   4.50000000    4.91666221    9.26%
19:   4.50000000    4.92487984    9.44%
Solution saved to mesh.vtk
```

Numerical solution ($a = 15$, $\omega = 1$, 290620 nodes):

E	1	2-4	5-10	11-
theory	1.5	2.5	3.5	4.5
FEM	1.522	2.535	3.554	4.578
		2.536	3.555	4.579
		2.536	3.555	4.579
			3.555	...
			3.556	
			3.556	

For better imagination, we plotted the eigenvector 5 from several sides and then eigenvector 10, see the figs 4.10 – 4.13.

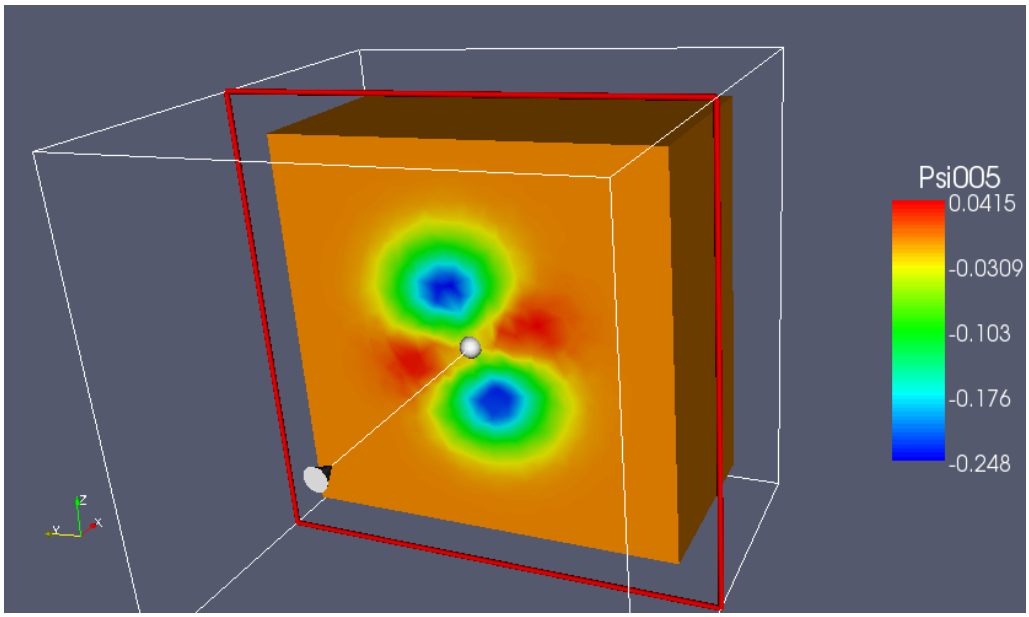


Figure 4.10: 3D oscillator: eigenvector 5, cut plane

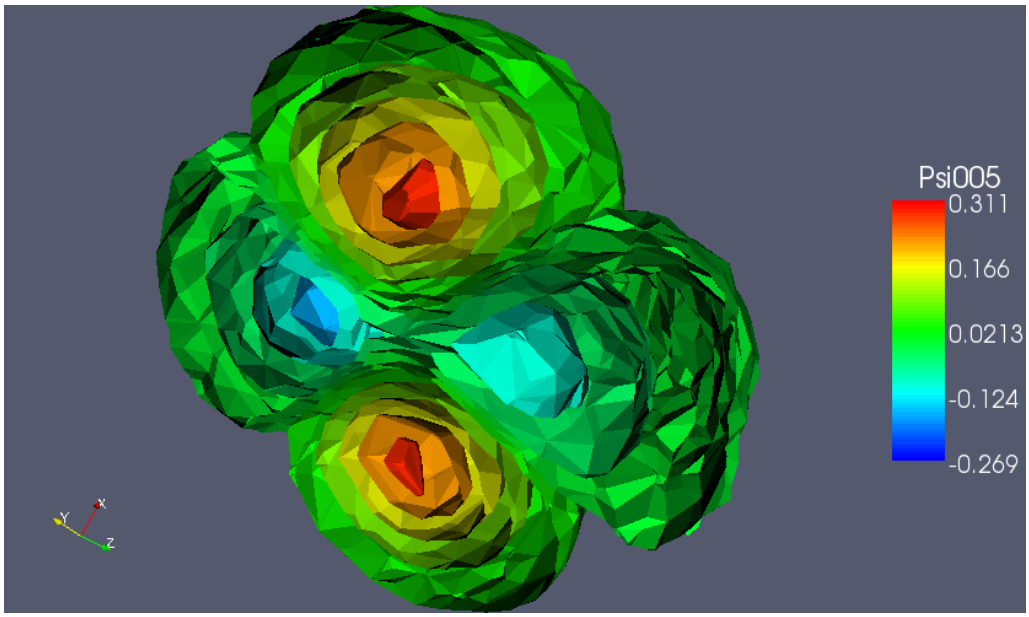


Figure 4.11: 3D oscillator: eigenvector 5, cut plane and contour

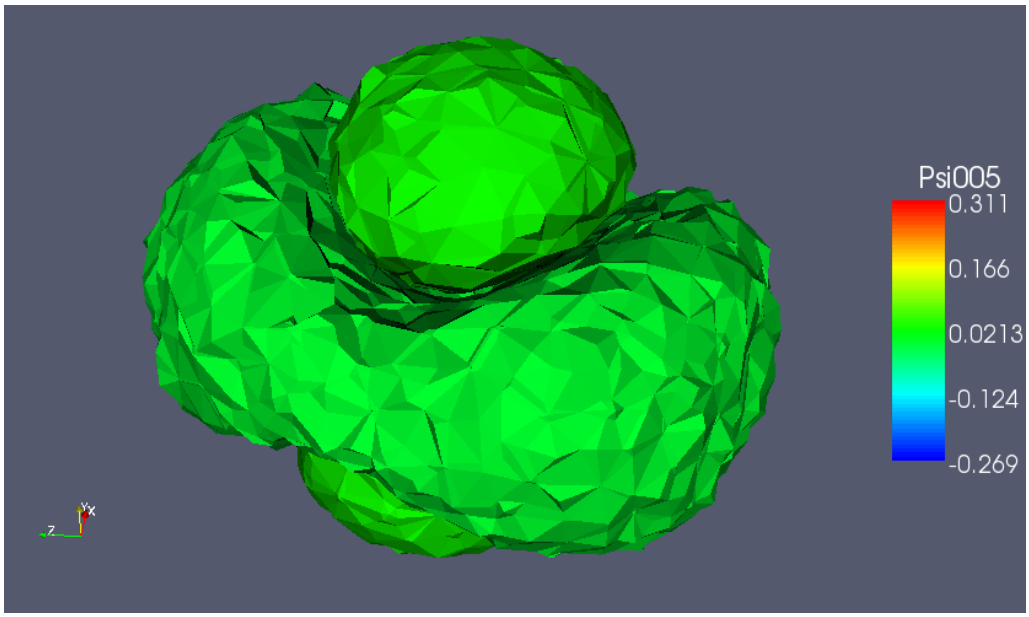


Figure 4.12: 3D oscillator: eigenvector 5, contour

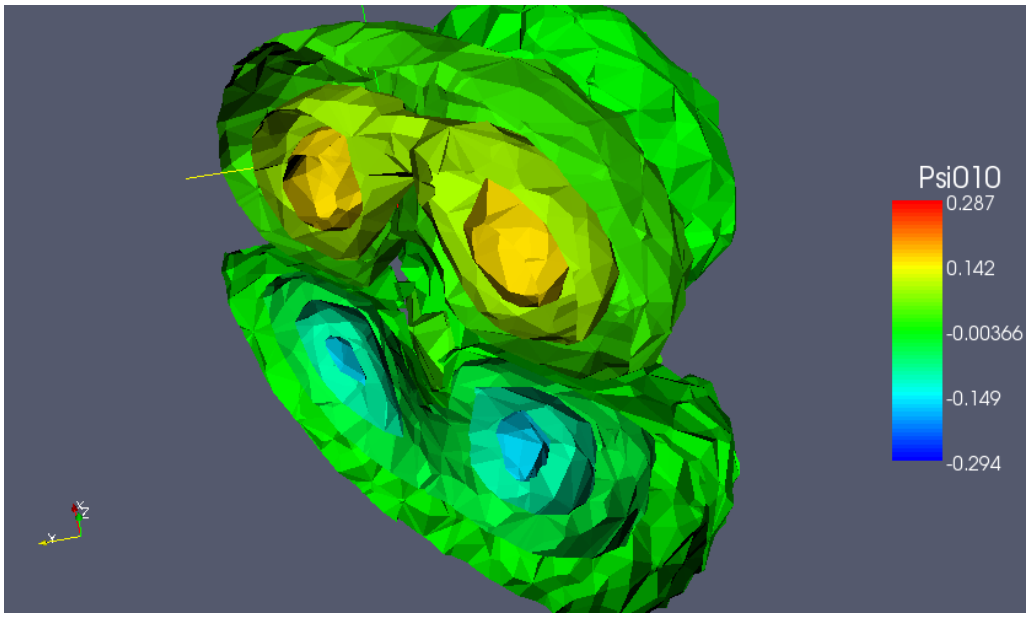


Figure 4.13: 3D oscillator: eigenvector 10, cut plane and contour

4.1.9 3D: Hydrogen Atom

$$V(r) = \begin{cases} -\frac{Z^2}{r}, & \text{inside the box} \quad a \times a \times a \\ \infty, & \text{outside} \end{cases}$$

Analytic solution in the limit $a \to \infty$:

$$E_n = -\frac{Z^2}{2n^2}$$

where $n = 1, 2, 3, \ldots$. Degeneracy is n^2, so for $Z = 1$: $E_1 = -\frac{1}{2} = -0.5$, $E_2 = -\frac{1}{8} = -0.125$, $E_3 = -\frac{1}{18} = -0.055$, $E_4 = -\frac{1}{32} = -0.031$.

```
$ ./schroedinger.py --hydrogen
Dimension: 3
[...]
Energies:
n       exact          FEM
0:  -0.50000000   -0.13468961
1:  -0.12500000    0.13909268
2:  -0.12500000    0.13934116
3:  -0.12500000    0.13939501
4:  -0.12500000    0.26835117
Solution saved to mesh.vtk
```

As you can see above, our mesh is not sufficient to get precise results, so we used a refined mesh with $a = 15$ and 160000 nodes and we got more precise results:

E	1	2-5	6-14	15-
theory	-0.5	-0.125	-0.055	-0.031
FEM	-0.481	-0.118	-0.006	$\cdots$

4.1.10 2D: nonsymmetric potential I

In this example we use a potential from two nuclei positioned at $(-5, 0)$ and $(5, 0)$. This is a nonsymmetric problem, thus one cannot use the usual way to reduce the Schrödinger equation to radial and angular parts. A general partial differential equations solver (in our case FEM) has to be used. See the fig 4.14 for the output of the lowest 10 eigenvectors.

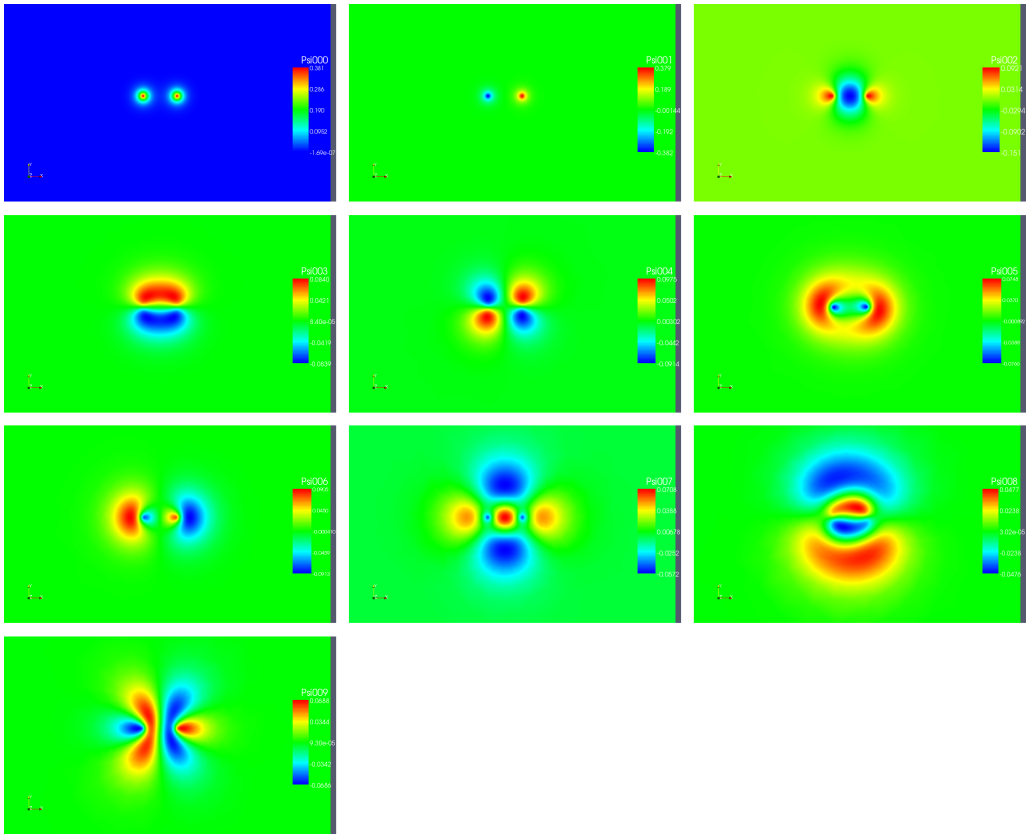


Figure 4.14: 2D nonsymmetric potential I: eigenvectors 0 to 9

4.1.11 2D: nonsymmetric potential II

In this example we use a potential from three nuclei positioned at $(-5, 0)$, $(5, 0)$ and $(0, 5)$. This is a similar problem to the previous example, but this time we plotted the solution using colors in a 2D plane (fig. 4.15) and also as a 3D surface (fig 4.16), so that the reader can get a better feeling about the solutions.

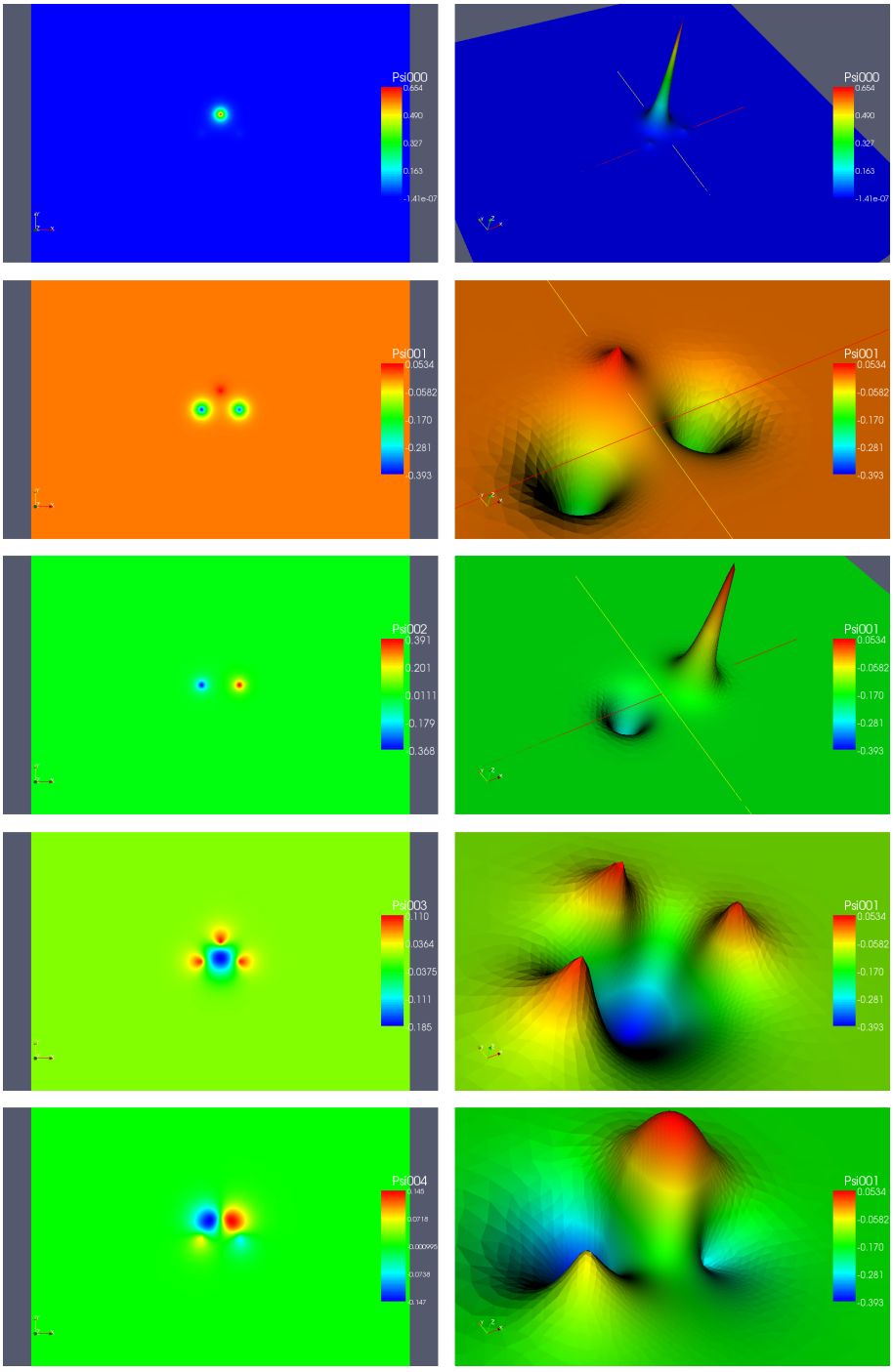


Figure 4.15: 2D nonsymmetric potential II: eigenvectors 0 to 4, plot in 2d on the left, surface plot on the right

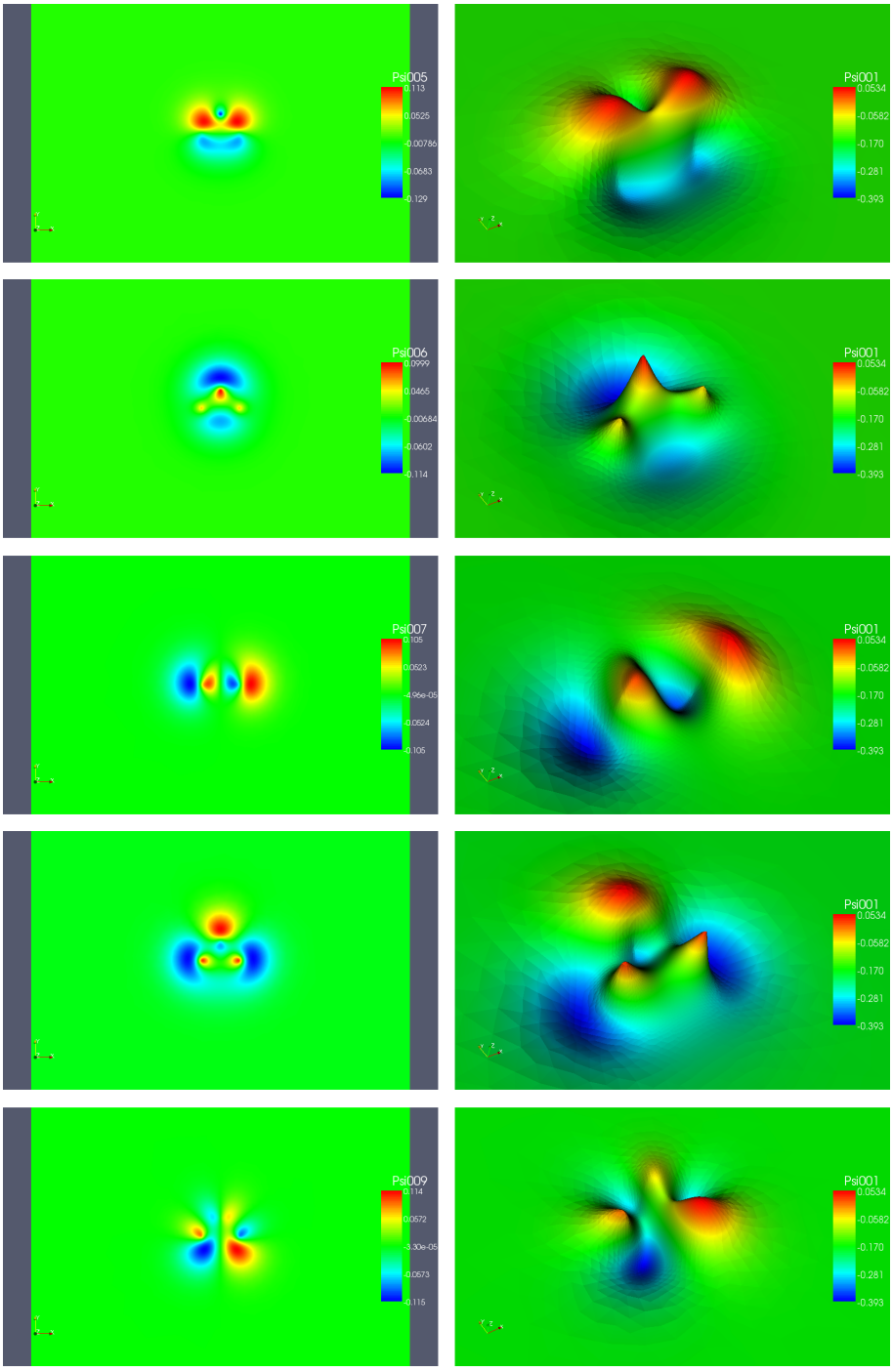


Figure 4.16: 2D nonsymmetric potential II: eigenvectors 5 to 9, plot in 2d on the left, surface plot on the right

4.1.12 2D: nonsymmetric potential III

Finally we calculate the Schrödinger equation for 20 atom nuclei uniformly positioned on a circle. We plotted several eigenvectors, again in 2D and as a 3D surface for better imagination, see the fig 4.17.

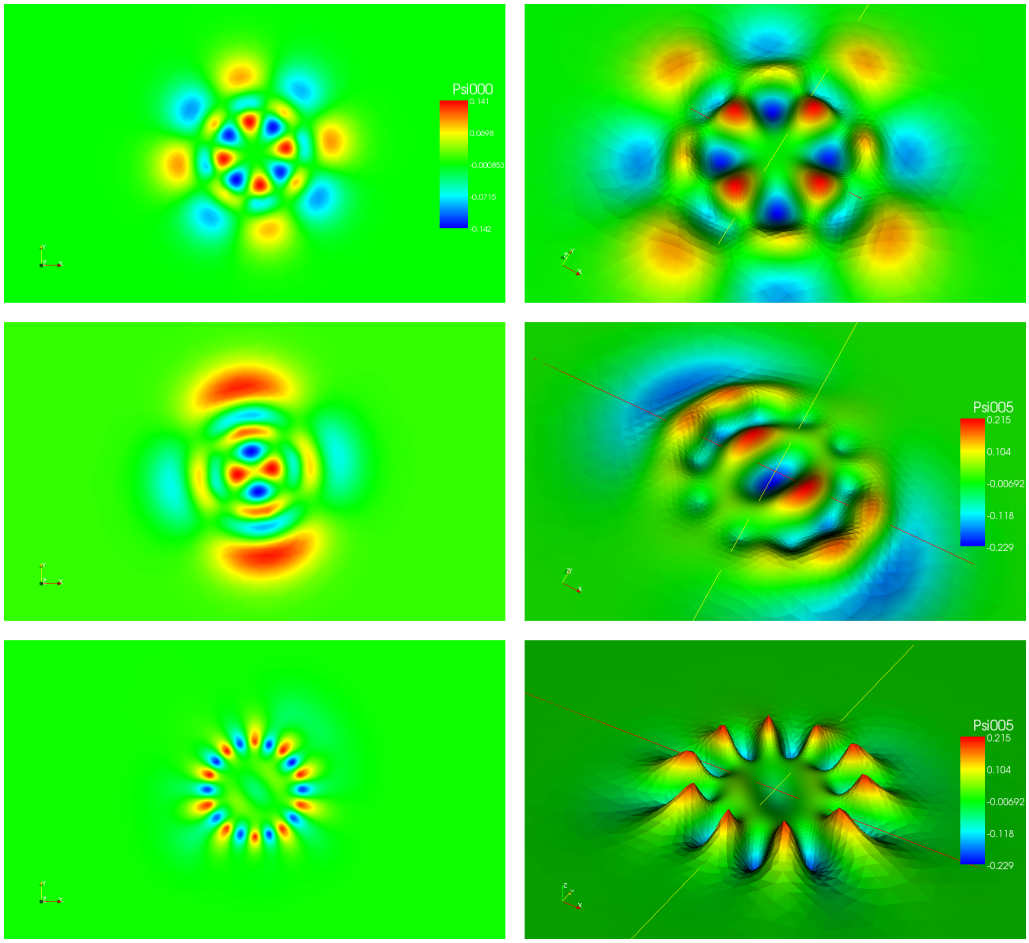


Figure 4.17: 2D nonsymmetric potential III: eigenvectors 0, 4 and 5, plot in 2D on the left, surface plot on the right

4.2 Density Functional Theory, Spherically Symmetric Solution

In this section we show solutions for the relativistic and non-relativistic atoms, in particular Lead and Boron. This is necessary to test the density functional theory self-consistency cycle code and also that all the parts of the program are working correctly. We also need some referential data to compare with. We use [12], that contains both relativistic and non-relativistic Kohn-Sham eigenvalues for every atom. See the section 3.1 for thorough introduction and derivation of all needed equations, so let's just briefly recapitulate: the potential for atoms is spherically symmetric, so the Kohn-Sham equations effectively become radial Schrödinger or Dirac equations, that is solved on a radial grid using the methods described in the section 3.1.

4.2.1 Pb: LDA

For nonrelativistic Lead, we need to solve the Schrödinger equation for 82 electrons. Here is a result of our program:

```
0:  |F(x)|=32381.03274296
1:  |F(x)|=5347.06647460
2:  |F(x)|=2283.82989218
3:  |F(x)|=148.26174998
4:  |F(x)|=120.78098800
5:  |F(x)|=84.68248231
6:  |F(x)|=11.30008638
7:  |F(x)|=3.65004163
8:  |F(x)|=3.12545615
9:  |F(x)|=1.44414848
10:  |F(x)|=0.32879840
11:  |F(x)|=0.10891716
12:  |F(x)|=0.03456829
13:  |F(x)|=0.01240870
14:  |F(x)|=0.00774382
15:  |F(x)|=0.00302906
16:  |F(x)|=0.00081825
17:  |F(x)|=0.00026270
18:  |F(x)|=0.00007814
19:  |F(x)|=0.00003516
```

```
1s( 2) j=l+1/2: -2901.078061
2s( 2) j=l+1/2: -488.8433352
2p( 6) j=l+1/2: -470.8777849
3s( 2) j=l+1/2: -116.526852
3p( 6) j=l+1/2: -107.950391
3d(10) j=l+1/2: -91.88992429
4s( 2) j=l+1/2: -25.75333021
4p( 6) j=l+1/2: -21.99056413
4d(10) j=l+1/2: -15.03002657
4f(14) j=l+1/2: -5.592531664
5s( 2) j=l+1/2: -4.206797624
5p( 6) j=l+1/2: -2.941656967
5d(10) j=l+1/2: -0.9023926829
6s( 2) j=l+1/2: -0.3571868295
6p( 2) j=l+1/2: -0.1418313263
```

The $|F(x)|$ shows the norm of the residual in the non-linear iteration, in this particular case (and all the other examples below), we use the mixing scheme described by the equation (3.28) with $\alpha = 0.3$ and $\alpha_{max} = 1.0$ (see the section 3.2.6 for more details). The rest of the output are individual Kohn-Sham energies.

The NIST reference calculation[12] has precision of 6 decimal places and the above calculation agrees with it to every decimal digit. This shows, that our implementation of the density functional theory cycle, Poisson's equation, LDA and other things is correct.

4.2.2 Pb: RLDA

For a relativistic Lead we need to solve a radial Dirac equation in the framework of DFT, together with a relativistic exchange and correlation potential.

```
0: |F(x)|=41056.67822797
1: |F(x)|=7926.04357205
2: |F(x)|=2858.46358638
3: |F(x)|=598.63802038
4: |F(x)|=268.50208068
5: |F(x)|=30.05744371
6: |F(x)|=27.56292000
7: |F(x)|=11.22084649
8: |F(x)|=4.78645898
```

```
9:  |F(x)|=0.53300950
10:  |F(x)|=0.13956963
11:  |F(x)|=0.07821891
12:  |F(x)|=0.06505839
13:  |F(x)|=0.02021479
14:  |F(x)|=0.00240256
15:  |F(x)|=0.00132181
16:  |F(x)|=0.00079552
17:  |F(x)|=0.00018579
18:  |F(x)|=0.00000838
19:  |F(x)|=0.00000584
1s( 2) j=l+1/2:  -3209.51946
2s( 2) j=l+1/2:  -574.1825655
2p( 6) j=l-1/2:  -551.7234408
2p( 6) j=l+1/2:  -472.3716103
3s( 2) j=l+1/2:  -137.8642241
3p( 6) j=l-1/2:  -127.6789451
3p( 6) j=l+1/2:  -109.9540395
3d(10) j=l-1/2:  -93.15817605
3d(10) j=l+1/2:  -89.36399096
4s( 2) j=l+1/2:  -31.15015728
4p( 6) j=l-1/2:  -26.73281564
4p( 6) j=l+1/2:  -22.38230707
4d(10) j=l-1/2:  -15.1647618
4d(10) j=l+1/2:  -14.3484973
5s( 2) j=l+1/2:  -5.225938506
4f(14) j=l-1/2:  -4.960490099
4f(14) j=l+1/2:  -4.775660273
5p( 6) j=l-1/2:  -3.710458943
5p( 6) j=l+1/2:  -2.889127431
5d(10) j=l-1/2:  -0.8020049565
5d(10) j=l+1/2:  -0.7070299184
6s( 2) j=l+1/2:  -0.4209603386
6p( 2) j=l-1/2:  -0.1549640727
```

Most of the states above agree with the NIST reference calculation[12] to one decimal place after the floating point (for example the 6p state differs by 0.02 Hartees, the 1p state by 0.04 Hartrees), the difference probably being caused by a different implementation of the exchange and correlation

potential approximation (in spite of the fact that both our code and NIST uses the XC potential proposed by MacDonald and Vosko). In particular, we implemented the equation (3.25), see also the discussion leading to it.

However, if you compare for example the energies of the 1s state in the relativistic (-3209.51946 Hartrees) an non relativistic calculation (-2901.078061 Hartrees), the difference is about 308.44 Hartrees and both energies agree with NIST (relativistic up to 0.04 Hartrees, nonrelativistic up to 10^{-6} Hartrees, e.g. all decimal places showed at NIST), so it is clear, that whatever causes the small difference, it is not a major problem.

4.2.3 B: LDA

We calculate nonrelativistic Boron.

```
1:   |F(x)|=467.33470427
2:   |F(x)|=39.46088238
3:   |F(x)|=5.59717305
4:   |F(x)|=3.09300726
5:   |F(x)|=2.04909614
6:   |F(x)|=0.09754169
7:   |F(x)|=0.06773803
8:   |F(x)|=0.04587578
9:   |F(x)|=0.00592044
10:   |F(x)|=0.00382678
11:   |F(x)|=0.00232014
12:   |F(x)|=0.00005561
13:   |F(x)|=0.00002714
14:   |F(x)|=0.00001809
15:   |F(x)|=0.00000042
16:   |F(x)|=0.00000023
17:   |F(x)|=0.00000014
18:   |F(x)|=0.00000001
19:   |F(x)|=0.00000000
20:   |F(x)|=0.00000000
1s(  2)  j=l+1/2:  -6.564347081
2s(  2)  j=l+1/2:  -0.3447010093
2p(  1)  j=l+1/2:  -0.1366031499
```

Agrees with NIST[12] to all decimal places.

4.2.4 B: RLDA

Finally we calculate a relativistic Boron.

```
0:  |F(x)|=485.06815695
1:  |F(x)|=103.13739852
2:  |F(x)|=34.94893789
3:  |F(x)|=18.15071235
4:  |F(x)|=1.19766447
5:  |F(x)|=0.10974802
6:  |F(x)|=0.07628667
7:  |F(x)|=0.02450426
8:  |F(x)|=0.00430650
9:  |F(x)|=0.00193614
10:  |F(x)|=0.00042938
11:  |F(x)|=0.00010306
12:  |F(x)|=0.00002872
13:  |F(x)|=0.00001073
14:  |F(x)|=0.00000217
15:  |F(x)|=0.00000090
16:  |F(x)|=0.00000028
17:  |F(x)|=0.00000015
18:  |F(x)|=0.00000002
19:  |F(x)|=0.00000001
1s(  2)  j=l+1/2:  -6.56282977
2s(  2)  j=l+1/2:  -0.3447247582
2p(  1)  j=l-1/2:  -0.1366103284
```

Agrees with NIST[12] to 3 decimal places after the decimal dot, see the comments at the relativistic Lead above for a discussion why it is different.

Chapter 5

Conclusion

In this thesis, we introduce the ingredients that are necessary for the new method for electronic structure calculations of non-periodic systems that should be based on density functional theory, ab-initio pseudopotentials and finite element method. We reviewed all the theory that we need, in particular density functional theory, radial Dirac and Schrödinger equations, Kohn-Sham equations, pseudopotentials and their separable form, and finite elements.

We then show thorough examples of results of testing some of the ingredients — i.e. the pieces of the code dedicated to solve the partial tasks — that we can calculate within the code developed so far: one electron Schrödinger equation solved by finite elements in 2D and 3D, for symmetric and non-symmetric potentials, and a selfconsistency engine within density functional theory for radial Dirac and Schrödinger equations.

All of that are the essential ingredients for our ultimate goal mentioned above and we show that this approach seems to be viable – all of the ingredients are there and working.

In the future work, what needs to be done is to take our DFT self-consistency engine tested on the radial Dirac / Schrödinger problem and apply our finite element solver for Kohn-Sham equations with separable pseudopotentials.

Appendix

5.1 Polar and Spherical Coordinates

Polar coordinates (radial, azimuth) (r, ϕ) are defined by

$$\begin{aligned} x &= r\cos\phi \\ y &= r\sin\phi \end{aligned}$$

Spherical coordinates (radial, zenith, azimuth) (ρ, θ, ϕ):

$$\begin{aligned} x &= \rho\sin\theta\cos\phi \\ y &= \rho\sin\theta\sin\phi \\ z &= \rho\cos\theta \end{aligned}$$

Note: this meaning of (θ, ϕ) is mostly used in the USA and in many books. In Europe people usualy use different symbols, like (ϕ, θ), (ϑ, φ) and others.

5.2 Delta Function

Delta function $\delta(x)$ is defined such that this relation holds:

$$\int f(x)\delta(x-t)\mathrm{d}x = f(t) \tag{5.1}$$

No such function exists, but one can find many sequences "converging" to a delta function:

$$\lim_{\alpha\to\infty} \delta_\alpha(x) = \delta(x) \tag{5.2}$$

more precisely:

$$\lim_{\alpha\to\infty}\int f(x)\delta_\alpha(x)\mathrm{d}x = \int f(x)\lim_{\alpha\to\infty}\delta_\alpha(x)\mathrm{d}x = f(t) \tag{5.3}$$

one example of such a sequence is:

$$\delta_\alpha(x) = \frac{1}{\pi x}\sin(\alpha x)$$

It's clear that (5.3) holds for any well behaved function $f(x)$. Mathematicians like to say it's incorrect to use such a notation when in fact the integral (5.1) doesn't "exist", but they are wrong, because it is not important if something "exist" or not, but rather if it is clear what we mean by our notation: (5.1) is a shorthand for (5.3) and (5.2) gets a mathematically rigorous meaning when you integrate both sides and use (5.1) to arrive at (5.3). Thus one uses the relations (5.1), (5.2), (5.3) to derive all properties of the delta function.

Let's give an example. Let $\hat{\mathbf{r}}$ be the unit vector in 3D and we can label it using spherical coordinates $\hat{\mathbf{r}} = \hat{\mathbf{r}}(\theta,\phi)$. We can also express it in cartesian coordinates as $\hat{\mathbf{r}}(\theta,\phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$.

$$f(\hat{\mathbf{r}}') = \int \delta(\hat{\mathbf{r}}-\hat{\mathbf{r}}')f(\hat{\mathbf{r}})\,\mathrm{d}\hat{\mathbf{r}} \tag{5.4}$$

Expressing $f(\hat{\mathbf{r}}) = f(\theta,\phi)$ as a function of θ and ϕ we have

$$f(\theta',\phi') = \int \delta(\theta-\theta')\delta(\phi-\phi')f(\theta,\phi)\,\mathrm{d}\theta\mathrm{d}\phi \tag{5.5}$$

Expressing (5.4) in spherical coordinates we get

$$f(\theta',\phi') = \int \delta(\hat{\mathbf{r}}-\hat{\mathbf{r}}')f(\theta,\phi)\sin\theta\,\mathrm{d}\theta\mathrm{d}\phi$$

and comparing to (5.5) we finally get

$$\delta(\hat{\mathbf{r}}-\hat{\mathbf{r}}') = \frac{1}{\sin\theta}\delta(\theta-\theta')\delta(\phi-\phi')$$

In exactly the same manner we get

$$\delta(\mathbf{r}-\mathbf{r}') = \delta(\hat{\mathbf{r}}-\hat{\mathbf{r}}')\frac{\delta(\rho-\rho')}{\rho^2}$$

See also (5.6) for an example of how to deal with more complex expressions involving the delta function like $\delta^2(x)$.

5.3 Variations and functional derivatives

Functional derivatives are a common source of confusion and especially the notation. The reason is similar to the delta function — the definition is operational, i.e. it tells you what operations you need to do to get a mathematically precise formula. The notation below is commonly used in physics and in our opinion it is perfectly precise and exact, but some mathematicians may not like it.

Let's have $\mathbf{x} = (x_1, x_2, \ldots, x_N)$. The function $f(\mathbf{x})$ assigns a number to each $\mathbf{x}$. We define a differential of f as

$$\mathrm{d}f \equiv \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon \mathrm{h})\right|_{\varepsilon=0} = \lim_{\varepsilon \to 0} \frac{f(\mathbf{x} + \varepsilon \mathrm{h}) - f(\mathbf{x})}{\varepsilon} = \mathrm{a} \cdot \mathrm{h}$$

The last equality follows from the fact, that $\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon \mathrm{h})\right|_{\varepsilon=0}$ is a linear function of h. We define $\frac{\partial f}{\partial x_i}$ as

$$\mathrm{a} \equiv \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_N}\right)$$

This also gives a formula for computing $\frac{\partial f}{\partial x_i}$: we set $h_j = \delta_{ij} h_i$ and

$$\frac{\partial f}{\partial x_i} = a_i = \mathrm{a} \cdot \mathrm{h} = \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon(0, 0, \ldots, 1, \ldots, 0))\right|_{\varepsilon=0} =$$

$$= \lim_{\varepsilon \to 0} \frac{f(x_1, x_2, \ldots, x_i + \varepsilon, \ldots, x_N) - f(x_1, x_2, \ldots, x_i, \ldots, x_N)}{\varepsilon}$$

But this is just the way the partial derivative is usually defined. Every variable can be treated as a function (very simple one):

$$x_i = g(x_1, \ldots, x_N) = \delta_{ij} x_j$$

and so we define

$$\mathrm{d}x_i \equiv \mathrm{d}g = \mathrm{d}(\delta_{ij} x_j) = h_i$$

and thus we write $h_i = \mathrm{d}x_i$ and $\mathrm{h} = \mathrm{d}\mathbf{x}$ and

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}x_i} \mathrm{d}x_i$$

So $\mathrm{d}\mathbf{x}$ has two meanings — it's either $\mathrm{h} = \mathbf{x} - \mathbf{x}_0$ (a finite change in the independent variable $\mathbf{x}$) or a differential, depending on the context. Even mathematicians use this notation.

Functional $F[f]$ assigns a number to each function $f(x)$. The variation is defined as

$$\delta F[f] \equiv \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F[f+\varepsilon h]\bigg|_{\varepsilon=0} = \lim_{\epsilon\to 0} \frac{F[f+\epsilon h]-F[f]}{\epsilon} = \int a(x)h(x)\mathrm{d}x$$

We define $\frac{\delta F}{\delta f(x)}$ as

$$a(x) \equiv \frac{\delta F}{\delta f(x)}$$

This also gives a formula for computing $\frac{\delta F}{\delta f(x)}$: we set $h(y) = \delta(x-y)$ and

$$\frac{\delta F}{\delta f(x)} = a(x) = \int a(y)\delta(x-y)\mathrm{d}y = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F[f(y)+\varepsilon\delta(x-y)]\bigg|_{\varepsilon=0} =$$

$$= \lim_{\varepsilon\to 0} \frac{F[f(y)+\varepsilon\delta(x-y)]-F[f(y)]}{\varepsilon}$$

Every function can be treated as a functional (although a very simple one):

$$f(x) = G[f] = \int f(y)\delta(x-y)\mathrm{d}y$$

and so we define

$$\delta f \equiv \delta G[f] = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} G[f(x)+\varepsilon h(x)]\bigg|_{\varepsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(x)+\varepsilon h(x))\bigg|_{\varepsilon=0} = h(x)$$

thus we write $h = \delta f$ and

$$\delta F[f] = \int \frac{\delta F}{\delta f(x)}\delta f(x)\mathrm{d}x$$

so δf have two meanings — it's either $h(x) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(x)+\varepsilon h(x))\big|_{\varepsilon=0}$ (a finite change in the function f) or a variation of a functional, depending on the context. Mathematicians never write δf in the meaning of $h(x)$, they always write the latter, but it's ridiculous, because it is completely analogous to $\mathrm{d}\mathbf{x}$.

The correspondence between the finite and infinite dimensional case can

be summarized as:

$$\begin{array}{rcl} f(x_i) & \Longleftrightarrow & F[f] \\ \mathrm{d}f = 0 & \Longleftrightarrow & \delta F = 0 \\ \dfrac{\partial f}{\partial x_i} = 0 & \Longleftrightarrow & \dfrac{\delta F}{\delta f(x)} = 0 \\ f & \Longleftrightarrow & F \\ x_i & \Longleftrightarrow & f(x) \\ x & \Longleftrightarrow & f \\ i & \Longleftrightarrow & x \end{array}$$

More generally, δ-variation can by applied to any function g which contains the function $f(x)$ being varied, you just need to replace f by $f + \epsilon h$ and apply $\frac{\mathrm{d}}{\mathrm{d}\epsilon}$ to the whole g, for example (here $g = \partial_\mu \phi$ and $f = \phi$):

$$\delta \partial_\mu \phi = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \partial_\mu (\phi + \varepsilon h)\bigg|_{\varepsilon=0} = \partial_\mu \frac{\mathrm{d}}{\mathrm{d}\varepsilon} (\phi + \varepsilon h)\bigg|_{\varepsilon=0} = \partial_\mu \delta\phi$$

This notation allows us a very convinient computation, as shown in the following examples. First, when computing a variation of some integral, when can interchange δ and $\int$:

$$F[f] = \int K(x) f(x) \mathrm{d}x$$

$$\delta F = \delta \int K(x) f(x) \mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int K(x)(f + \varepsilon h) \mathrm{d}x\bigg|_{\varepsilon=0} = \int \frac{\mathrm{d}}{\mathrm{d}\varepsilon} (K(x)(f + \varepsilon h)) \mathrm{d}x\bigg|_{\varepsilon=0} =$$

$$= \int \delta(K(x) f(x)) \mathrm{d}x$$

In the expression $\delta(K(x)f(x))$ we must understand from the context if we are treating it as a functional of f or K. In our case it's a functional of f, so we have $\delta(Kf) = K\delta f$.

A few more examples:

$$\frac{\delta}{\delta f(t)} \int \mathrm{d}t' f(t') g(t') = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int \mathrm{d}t' (f(t') + \varepsilon \delta(t - t')) g(t')\bigg|_{\varepsilon=0} = g(t)$$

$$\frac{\delta f(t')}{\delta f(t)} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(t') + \varepsilon\delta(t-t'))\bigg|_{\varepsilon=0} = \delta(t-t')$$

$$\frac{\delta f(t_1)f(t_2)}{\delta f(t)} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(t_1) + \varepsilon\delta(t-t_1))(f(t_2) + \varepsilon\delta(t-t_2))\bigg|_{\varepsilon=0} = \delta(t-t_1)f(t_2)+f(t_1)\delta(t-t_2)$$

$$\frac{\delta}{\delta f(t)}\tfrac{1}{2}\int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)f(t_1)f(t_2) = \tfrac{1}{2}\int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)\frac{\delta f(t_1)f(t_2)}{\delta f(t)} =$$

$$= \tfrac{1}{2}\left(\int \mathrm{d}t_1 K(t_1,t)f(t_1) + \int \mathrm{d}t_2 K(t,t_2)f(t_2)\right) = \int \mathrm{d}t_2 K(t,t_2)f(t_2)$$

The last equality follows from $K(t_1,t_2) = K(t_2,t_1)$ (any antisymmetrical part of a K would not contribute to the symmetrical integration).

$$\frac{\delta}{\delta f(t)}\int f^3(x)\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int (f(x) + \varepsilon\delta(x-t))^3\mathrm{d}x\bigg|_{\varepsilon=0} =$$

$$= \int 3(f(x) + \varepsilon\delta(x-t))^2\delta(x-t)\mathrm{d}x\bigg|_{\varepsilon=0} = \int 3f^2(x)\delta(x-t)\mathrm{d}x = 3f^2(t)$$

Some mathematicians would say the above calculation is incorrect, because $\delta^2(x-t)$ is undefined. But that's not true, because in case of such problems the above notation automatically implies working with some sequence $\delta_\alpha(x) \to \delta(x)$ (for example $\delta_\alpha(x) = \frac{1}{\pi x}\sin(\alpha x)$) and taking the limit $\alpha \to \infty$:

$$\frac{\delta}{\delta f(t)}\int f^3(x)\mathrm{d}x = \lim_{\alpha\to\infty}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int (f(x) + \varepsilon\delta_\alpha(x-t))^3\mathrm{d}x\bigg|_{\varepsilon=0} =$$

$$= \lim_{\alpha\to\infty}\int 3(f(x) + \varepsilon\delta_\alpha(x-t))^2\delta_\alpha(x-t)\mathrm{d}x\bigg|_{\varepsilon=0} = \lim_{\alpha\to\infty}\int 3f^2(x)\delta_\alpha(x-t)\mathrm{d}x =$$

$$= \int 3f^2(x)\lim_{\alpha\to\infty}\delta_\alpha(x-t)\mathrm{d}x = \int 3f^2(x)\delta(x-t)\mathrm{d}x = 3f^2(t) \qquad (5.6)$$

As you can see, we got the same result, with the same rigor, but using an obfuscating notation. That's why such obvious manipulations with δ_α are tacitly implied.

5.4 Spherical Harmonics

Are defined by

$$Y_{lm}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos\theta)\, e^{im\phi}$$

where P_l^m are associated Legendre polynomials defined by

$$P_l^m(x) = (-1)^m(1-x^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}x^m}P_l(x)$$

and P_l are Legendre polynomials defined by the formula

$$P_l(x) = \frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}x^l}[(x^2-1)^l]$$

they also obey the completeness relation

$$\sum_{l=0}^{\infty}\frac{2l+1}{2}P_l(x')P_l(x) = \delta(x-x') \tag{5.7}$$

The spherical harmonics are ortonormal:

$$\int Y_{lm}\, Y^*_{l'm'}\,\mathrm{d}\Omega = \int_0^{2\pi}\int_0^{\pi} Y_{lm}(\theta,\phi)\, Y^*_{l'm'}(\theta,\phi)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi = \delta_{mm'}\delta_{ll'} \tag{5.8}$$

and complete (both in the l-subspace and the whole space):

$$\sum_{m=-l}^{l}|Y_{lm}(\theta,\phi)|^2 = \frac{2l+1}{4\pi} \tag{5.9}$$

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\theta,\phi)Y^*_{lm}(\theta',\phi') = \frac{1}{\sin\theta}\delta(\theta-\theta')\delta(\phi-\phi') = \delta(\hat{\mathbf{r}}-\hat{\mathbf{r}}') \tag{5.10}$$

The relation (5.9) is a special case of an addition theorem for spherical harmonics

$$\sum_{m=-l}^{l} Y_{lm}(\theta,\phi)Y^*_{lm}(\theta',\phi') = \frac{4\pi}{2l+1}P_l(\cos\gamma) \tag{5.11}$$

where γ is the angle between the unit vectors given by $\hat{\mathbf{r}} = (\theta,\phi)$ and $\hat{\mathbf{r}}' = (\theta',\phi')$:

$$\cos\gamma = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi-\phi') = \hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'$$

5.5 Dirac Notation

The Dirac notation allows a very compact and powerful way of writing equations that describe a function expansion into a basis, both discrete (e.g. a fourier series expansion) and continuous (e.g. a fourier transform) and related things. The notation is designed so that it is very easy to remember and it just guides you to write the correct equation.

Let's have a function $f(x)$. We define

$$\begin{aligned} \langle x|f\rangle &\equiv f(x) \\ \langle x'|f\rangle &\equiv f(x') \\ \langle x'|x\rangle &\equiv \delta(x'-x) \\ \int |x\rangle\langle x|\,\mathrm{d}x &\equiv \mathbb{1} \end{aligned}$$

The following equation

$$f(x') = \int \delta(x'-x)f(x)\mathrm{d}x$$

then becomes

$$\langle x'|f\rangle = \int \langle x'|x\rangle\langle x|f\rangle\,\mathrm{d}x$$

and thus we can interpret $|f\rangle$ as a vector, $|x\rangle$ as a basis and $\langle x|f\rangle$ as the coefficients in the basis expansion:

$$|f\rangle = \mathbb{1}\,|f\rangle = \int |x\rangle\langle x|\,\mathrm{d}x\,|f\rangle = \int |x\rangle\langle x|f\rangle\,\mathrm{d}x$$

That's all there is to it. Take the above rules as the operational definition of the Dirac notation. It's like with the delta function - written alone it doesn't have any meaning, but there are clear and non-ambiguous rules to convert any expression with δ to an expression which even mathematicians understand (i.e. integrating, applying test functions and using other relations to get rid of all δ symbols in the expression – but the result is usually much more complicated than the original formula). It's the same with the ket $|f\rangle$: written alone it doesn't have any meaning, but you can always use the above rules to get an expression that make sense to everyone (i.e. attaching any bra to the left and rewriting all brackets $\langle a|b\rangle$ with their equivalent

expressions) – but it will be more complex and harder to remember and – that is important – less general.

Now, let's look at the spherical harmonics:

$$Y_{lm}(\hat{\mathbf{r}}) \equiv \langle \hat{\mathbf{r}} | lm \rangle$$

on the unit sphere, we have

$$\int |\hat{\mathbf{r}}\rangle \langle \hat{\mathbf{r}}| \, \mathrm{d}\hat{\mathbf{r}} = \int |\hat{\mathbf{r}}\rangle \langle \hat{\mathbf{r}}| \, \mathrm{d}\Omega = \mathbb{1}$$

$$\delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') = \langle \hat{\mathbf{r}} | \hat{\mathbf{r}}' \rangle$$

thus

$$\int_0^{2\pi} \int_0^{\pi} Y_{lm}(\theta, \phi) \, Y^*_{l'm'}(\theta, \phi) \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi = \int \langle l'm' | \hat{\mathbf{r}} \rangle \langle \hat{\mathbf{r}} | lm \rangle \, \mathrm{d}\Omega = \langle l'm' | lm \rangle$$

and from (5.8) we get

$$\langle l'm' | lm \rangle = \delta_{mm'} \delta_{ll'}$$

now

$$\sum_{lm} Y_{lm}(\theta, \phi) Y^*_{lm}(\theta', \phi') = \sum_{lm} \langle \hat{\mathbf{r}} | lm \rangle \langle lm | \hat{\mathbf{r}}' \rangle$$

from (5.10) we get

$$\sum_{lm} \langle \hat{\mathbf{r}} | lm \rangle \langle lm | \hat{\mathbf{r}}' \rangle = \langle \hat{\mathbf{r}} | \hat{\mathbf{r}}' \rangle$$

so we have

$$\sum_{lm} |lm\rangle \langle lm| = \mathbb{1}$$

so $|lm\rangle$ forms an orthonormal basis. Any function defined on the sphere $f(\hat{\mathbf{r}})$ can be written using this basis:

$$f(\hat{\mathbf{r}}) = \langle \hat{\mathbf{r}} | f \rangle = \sum_{lm} \langle \hat{\mathbf{r}} | lm \rangle \langle lm | f \rangle = \sum_{lm} Y_{lm}(\hat{\mathbf{r}}) f_{lm}$$

where

$$f_{lm} = \langle lm | f \rangle = \int \langle lm | \hat{\mathbf{r}} \rangle \langle \hat{\mathbf{r}} | f \rangle \, \mathrm{d}\Omega = \int Y^*_{lm}(\hat{\mathbf{r}}) f(\hat{\mathbf{r}}) \mathrm{d}\Omega$$

If we have a function $f(\mathbf{r})$ in 3D, we can write it as a function of ρ and $\hat{\mathbf{r}}$ and expand only with respect to the variable $\hat{\mathbf{r}}$:

$$f(\mathbf{r}) = f(\rho\hat{\mathbf{r}}) \equiv g(\rho, \hat{\mathbf{r}}) = \sum_{lm} Y_{lm}(\hat{\mathbf{r}}) g_{lm}(\rho)$$

In Dirac notation we are doing the following: we decompose the space into the angular and radial part

$$|\mathbf{r}\rangle = |\hat{\mathbf{r}}\rangle \otimes |\rho\rangle \equiv |\hat{\mathbf{r}}\rangle\, |\rho\rangle$$

and write

$$f(\mathbf{r}) = \langle \mathbf{r}|f\rangle = \langle \hat{\mathbf{r}}|\, \langle \rho|f\rangle = \sum_{lm} Y_{lm}(\hat{\mathbf{r}})\, \langle lm|\, \langle \rho|f\rangle$$

where

$$\langle lm|\, \langle \rho|f\rangle = \int \langle lm|\hat{\mathbf{r}}\rangle\, \langle \hat{\mathbf{r}}|\, \langle \rho|f\rangle\, \mathrm{d}\Omega = \int Y^*_{lm}(\hat{\mathbf{r}}) f(\mathbf{r}) \mathrm{d}\Omega$$

Let's calculate $\langle \rho|\rho'\rangle$

$$\langle \mathbf{r}|\mathbf{r}'\rangle = \langle \hat{\mathbf{r}}|\, \langle \rho|\rho'\rangle\, |\hat{\mathbf{r}}'\rangle = \langle \hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle\, \langle \rho|\rho'\rangle$$

so

$$\langle \rho|\rho'\rangle = \frac{\langle \mathbf{r}|\mathbf{r}'\rangle}{\langle \hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle} = \frac{\delta(\rho - \rho')}{\rho^2}$$

We must stress that $|lm\rangle$ only acts in the $|\hat{\mathbf{r}}\rangle$ space (not the $|\rho\rangle$ space) which means that

$$\langle \mathbf{r}|lm\rangle = \langle \hat{\mathbf{r}}|\, \langle \rho|lm\rangle = \langle \hat{\mathbf{r}}|lm\rangle\, \langle \rho| = Y_{lm}(\hat{\mathbf{r}})\, \langle \rho|$$

and $V\, |lm\rangle$ leaves $V\, |\rho\rangle$ intact. Similarly,

$$\sum_{lm} |lm\rangle\, \langle lm| = \mathbb{1}$$

is a unity in the $|\hat{\mathbf{r}}\rangle$ space only (i.e. on the unit sphere).

Let's rewrite the equation (5.11):

$$\sum_{m} \langle \hat{\mathbf{r}}|lm\rangle\, \langle lm|\hat{\mathbf{r}}'\rangle = \frac{4\pi}{2l+1}\, \langle \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'|P_l\rangle$$

Using the completeness relation (5.7):

$$\sum_{l} \frac{2l+1}{2}\, \langle x'|P_l\rangle\, \langle P_l|x\rangle = \langle x'|x\rangle$$

$$\sum_{l} |P_l\rangle\, \frac{2l+1}{2}\, \langle P_l| = \mathbb{1}$$

we can now derive a very important formula true for every function $f(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')$:

$$f(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}') = \langle \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}' | f \rangle = \sum_l \langle \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}' | P_l \rangle \frac{2l+1}{2} \langle P_l | f \rangle = \sum_{lm} \langle \hat{\mathbf{r}} | lm \rangle \langle lm | \hat{\mathbf{r}}' \rangle \frac{(2l+1)^2}{8\pi} \langle P_l | f \rangle =$$

$$= \sum_{lm} \langle \hat{\mathbf{r}} | lm \rangle f_l \langle lm | \hat{\mathbf{r}}' \rangle$$

where

$$f_l = \frac{(2l+1)^2}{8\pi} \langle P_l | f \rangle = \frac{(2l+1)^2}{8\pi} \int_{-1}^{1} \langle P_l | x \rangle \langle x | f \rangle \, \mathrm{d}x = \frac{(2l+1)^2}{8\pi} \int_{-1}^{1} P_l(x) f(x) \mathrm{d}x$$

or written explicitly

$$f(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_{lm}(\hat{\mathbf{r}}) f_l Y_{lm}^*(\hat{\mathbf{r}}') \tag{5.12}$$

Bibliography

[1] G. B. Bachelet and M. Schluter. Relativistic norm-conserving pseudopotentials. *Phys. Rev. B*, 25(4), 1982.

[2] Peter E. Blochl. Generalized separable potentials for electronic-structure calculations. *Phys. Rev. B*, 41(5414), 1990.

[3] R. M. Dreizler and E.K.U.Gross. *Density Functional Theory*. Springer-Verlag, 1990.

[4] V. Eyert. A comparative study on methods for convergence acceleration of iterative vector sequences. *J. Comp. Phys.*, 124:271–285, 1996.

[5] Jiří Formánek. *Úvod do kvantové teorie*, volume I. Academia, Praha, 2004.

[6] V. Gavini, K. Bhattacharya, and M. Ortiz. Quasi-continuum orbital-free density-functional theory: A route to multi-million atom non-periodic dft calculation. *J. Mech. Phys. Solids*, 55:697–718, 2007.

[7] V. Gavini, J. Knap, K. Bhattacharya, and M. Ortiz. Non-periodic finite-element formulation of orbital-free density functional theory. *J. Mech. Phys. Solids*, 55:669–696, 2007.

[8] R. Geus. The jacobi–davidson algorithm for solving large sparse symmetric eigenvalue problems. Ph.D. Thesis No. 14734, ETH, Zurich, 2002.

[9] Christophe Geuzaine and Jean-François Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. `http://www.geuz.org/gmsh`, 1997-2008.

[10] A. Henderson. Paraview guide, a parallel visualization application. `http://paraview.org`, 2007.

[11] L. Kleinman and D. M. Bylander. Efficacious form for model pseudopotentials. *Phys. Rev. Lett.*, 48:1425–1428, 1982.

[12] S. Kotochigova, Z.H. Levine, E.L. Shirley, M.D. Stiles, and C.W. Clark. Local-density-functional calculations of the energy of atoms. *Phys. Rev. A*, 55:191–199, 1997. `http://physics.nist.gov/PhysRefData/DFTdata/contents.html`.

[13] A. H. MacDonald, W. E. Pickett, and D. D. Koelling. A linearised relativistic augmented-plane-wave method utilising approximate pure spin basis functions. *J. Phys. C: Solid St. Phys.*, 13:2675–83, 1980.

[14] Richard M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2005. ISBN 0-521-78285-6.

[15] J. E. Pask and P. A. Sterne. Finite element methods in ab initio electronic structure calculations. *Modelling Simul. Mater. Sci. Eng.*, 13, 2005.

[16] J. E. Pask and P. A. Sterne. Finite elements in ab initio electronic-structure calculations. *Handbook of Materials Modeling, S. Yip (ed.)*, page 423, 2005.

[17] W. E. Pickett. Pseudopotential methods in condensed matter applications. *Comp. Phys. Reports*, 9:115–198, 1989.

[18] G. P. Srivastava. Broyden's method for self-consistent field convergence acceleration. *J. Phys. A*, 17:L317–L321, 1984.

[19] Paul Strange. *Relativistic Quantum Mechanics*. Cambridge University Press, 1998.

[20] J. Vackář and A. Šimůnek. Adaptability and accuracy of all-electron pseudopotentials. *Phys. Rev. B*, 67(12712), 2003.

[21] D. Vanderbilt and S. G. Louie. Total energies of diamond (111) surface reconstructions by a linear combination of atomic orbitals method. *Phys. Rev. B*, 30(6118), 1984.

[22] J. Zabloudil, R. Hammerling, L. Szunyogh, and P. Weinberger. *Electron Scattering in Solid Matter*. Springer-Verlag Berlin, 2005. ISBN 3-540-22524-2.

[23] Ondřej Čertík. Numerical solution of the radial dirac equation in pseudopotential construction, 2007. arXiv:0710.3408 [physics.atom-ph].